

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**

Guilherme Klein da Silva Bitencourt

**UM MODELO DE CONFIANÇA BASEADO EM EMOÇÕES PARA
AGENTES**

Florianópolis

2014

Guilherme Klein da Silva Bitencourt

**UM MODELO DE CONFIANÇA BASEADO EM EMOÇÕES PARA
AGENTES**

Dissertação submetida ao Programa de
Pós-Graduação em Ciência da Computa-
ção da Universidade Federal de Santa Ca-
tarina para a obtenção do Grau de Mestre
em Ciência da Computação.

Orientador: Prof. Dr. Ricardo Azambuja
Silveira

Coorientadora: Profa. Dr^a. Jerusa Marchi

Florianópolis

2014

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Bitencourt, Guilherme Klein da Silva

Um modelo de confiança baseado em emoções para agentes /
Guilherme Klein da Silva Bitencourt ; orientador, Ricardo
Azambuja Silveira ; co-orientadora, Jerusa Marchi. -
Florianópolis, SC, 2014.

116 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Ciência da Computação.

Inclui referências

1. Ciência da Computação. 2. Agentes. 3. Confiança. 4.
Reputação. 5. Emoções. I. Silveira, Ricardo Azambuja. II.
Marchi, Jerusa. III. Universidade Federal de Santa
Catarina. Programa de Pós-Graduação em Ciência da Computação.
IV. Título.

Guilherme Klein da Silva Bitencourt

UM MODELO DE CONFIANÇA BASEADO EM EMOÇÕES PARA AGENTES

Esta Dissertação foi julgada adequada para a obtenção do Título de “Mestre em Ciência da Computação”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Santa Catarina.

Florianópolis, 04 de Fevereiro 2014.

Prof. Dr. Ronaldo dos Santos Mello
Coordenador

Banca Examinadora:

Prof. Dr. Ricardo Azambuja Silveira
Orientador

Profa. Dr^a. Jerusa Marchi
Coorientadora

Prof. Dr. Antônio Carlos da Rocha Costa
Universidade Federal do Rio Grande

Prof. Dr. Ricardo José Rabelo
Universidade Federal de Santa Catarina

Prof. Dr. Elder Rizzon dos Santos
Universidade Federal de Santa Catarina

Dedico esse trabalho aos meus pais, a minha namorada e a todos que de alguma forma me ajudaram durante esses três anos.

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me dado força e saúde para concluir esse trabalho. Aos meus pais, Visnoni de Bitencourt e Anita Maria Klein, pelo amor e carinho recebido ao longo da minha vida. Agradeço também ao meu orientador Ricardo Azambuja Silveira e a minha coorientadora Jerusa Marchi pela ajuda e compreensão ao longo dessa jornada.

RESUMO

Em sistemas multiagente, modelos de Confiança e Reputação vêm sendo propostos com o objetivo de auxiliar os agentes a selecionar seus parceiros de interação, visando aumentar o número de interações bem sucedidas. Assim como a confiança, emoções em agentes também são objetos de estudo tendo o objetivo de tornar as ações e reações dos agentes mais semelhantes a dos humanos. Essa dissertação apresenta um modelo de confiança baseado em emoções. Trata-se portanto de um modelo híbrido, que integra ao enfoque matemático característico dos modelos de confiança e reputação, um mecanismo de raciocínio simbólico, visando capturar a complexidade do raciocínio e flexibilizar as avaliações essencialmente quantitativas feitas pelos agentes através da introdução de elementos de natureza qualitativa ao cálculo da confiança. O modelo proposto incorpora oito emoções ao modelo de confiança Regret. Na parte inicial do trabalho é feita uma análise de modelos de confiança para agentes, do modelo OCC e de modelos de emoções para agentes. A parte principal do trabalho corresponde a proposta do modelo, que envolve a formalização das emoções e a modelagem das funções relativas às emoções. Ao final, dois cenários distintos são simulados para testar a funcionalidade do modelo.

Palavras-chave: Agentes, Confiança, Reputação, Emoções.

ABSTRACT

Trust and reputation have been proposed in the Multiagent Systems area as a way to assist agents to select good partners in order to improve interactions between them. As well as trust, agent emotions have been studied with the aim of making the actions and reactions of the agents more like those of humans. In this work, we present a trust model based on emotions. It is a hybrid mathematical and symbolical model which combines the approach of mathematical models of trust and reputation with a symbolic reasoning mechanism. It aims to capture the complexity of reasoning and easing the essentially quantitative assessments made by agents through the introduction of emotional elements in the trust evaluation. The proposed model incorporates eight emotions in Regret trust model. In the early part of the work an analysis of agent's trust models, OCC model and agent's emotions models is made. The main part of the work corresponds to the proposed model, including the formalization of emotions and modeling of functions related to emotions. At the end, two different scenarios are simulated to test the functionality of the model.

Keywords: Agents, Trust, Reputation, Emotions.

LISTA DE FIGURAS

Figura 1	Exemplo de gráfico de uma função beta (TEACY et al., 2006)	34
Figura 2	Especificações dos tipos de emoções do modelo OCC (ORTONY; CLORE; COLLINS, 1990) [tradução nossa]	54
Figura 3	Estrutura de emoções do modelo OCC (ORTONY; CLORE; COLLINS, 1990) [tradução nossa]	57
Figura 4	Hierarquia baseada em herança de emoções do modelo OCC modificado (STEUNEBRINK; DASTANI; MEYER, 2009) [tradução nossa]	69
Figura 5	Especificações dos tipos de emoções do modelo OCC modificado (STEUNEBRINK; DASTANI; MEYER, 2009) [tradução nossa] ..	70
Figura 6	Interface entre arquitetura EBDI e o modelo TrustE	75
Figura 7	Modelo TrustE	77
Figura 8	Visão Conceitual do Modelo TrustE	87

LISTA DE ABREVIATURAS E SIGLAS

IAD	Inteligência Artificial Distribuída	19
SMA	Sistemas Multiagente	19
EBDI	Emoções, Crenças, Desejos e Intenções	21
BDI	Crenças, Desejos e Intenções	25
IT	Confiança por interação	39
RT	Confiança baseada em papéis	41
WR	Reputação por testemunho	42
CR	Reputação Certificada	42
RI	Reputação Individual	46
RS	Reputação Social	47
EBDI	Emoções, Crenças, Desejos e Intenções	62
MEA	Memória de emoções dos agentes	73
EP	Emoções Positivas	76
EN	Emoções Negativas	76
IEE	Intensidade de Emoções disparadas por Eventos	77
IEA	Intensidade de Emoções disparadas por Ações	78

SUMÁRIO

1	INTRODUÇÃO	19
1.1	FORMULAÇÃO DO PROBLEMA	21
1.2	OBJETIVO	22
1.2.1	Objetivos Específicos	22
1.3	METODOLOGIA	22
1.4	ORGANIZAÇÃO DO TRABALHO	23
2	AGENTES E SMA	25
2.1	AGENTES	25
2.2	SISTEMAS MULTIAGENTE	26
2.2.1	Propriedades dos Ambientes MA	26
3	MODELOS DE CONFIANÇA EM SMA	29
3.1	MODELO DE MARSH	30
3.2	MODELO TRAVOS	31
3.2.1	Confiança e Confiabilidade	33
3.2.2	Reputação	35
3.2.3	Reputação Imprecisa	37
3.3	FIRE	39
3.3.1	Confiança por interação (IT)	39
3.3.2	Confiança baseada em papéis (RT)	41
3.3.3	Reputação por testemunho (WR)	42
3.3.4	Reputação Certificada (CR)	42
3.3.5	Confiança Final	43
3.4	REGRET	44
3.4.1	Resultado e Impressões	45
3.4.2	Reputação Individual	46
3.4.3	Reputação Social	47
3.4.4	Dimensão Ontológica	49
3.5	CONSIDERAÇÕES SOBRE OS MODELOS	50
3.5.1	Confiança <i>versus</i> Reputação	51
4	EMOÇÕES EM AGENTES INTELIGENTES	53
4.1	MODELO OCC	54
4.1.1	Tipos de emoções do modelo OCC	54
4.1.2	Hierarquia de emoções do OCC	56
4.1.3	Emoções baseadas em eventos	57
4.1.4	Emoções baseadas em agentes	58
4.1.5	Emoções baseadas em aspectos de objetos	59
4.1.6	Emoções compostas	59

4.2	MODELOS DE EMOÇÕES EM AGENTES	59
4.2.1	Modelo de Lorini	59
4.2.2	EBDI: Uma Arquitetura para Agentes Emocionais	62
4.2.3	Formalização de emoções OCC com lógica BDI	67
4.2.4	Modelo de Steunebrink	68
4.2.4.1	Nova hierarquia de emoções	68
4.2.4.2	Formalização das Emoções	69
4.2.4.3	Intensidade das Emoções	72
5	MODELO PROPOSTO	75
5.1	FORMALIZAÇÃO DAS EMOÇÕES NO TRUSTE.....	76
5.1.1	Emoções disparadas por eventos	77
5.1.2	Emoções disparadas por ações	78
5.2	INTENSIDADE DAS EMOÇÕES NO TRUSTE	79
5.2.1	Memória de emoções dos agentes.....	81
5.3	O MODELO TRUSTE	82
5.3.1	Funcionamento do TrustE	86
5.4	AVALIAÇÃO DO MODELO.....	88
5.4.1	Cenário 1	89
5.4.1.1	Simulação 1	92
5.4.2	Cenário 2	93
5.4.2.1	Simulação 1	96
6	CONSIDERAÇÕES FINAIS	99
6.1	TRABALHOS FUTUROS	100
	REFERÊNCIAS	101
	APÊNDICE A – Simulações Cenário 1.....	107
	APÊNDICE B – Simulações Cenário 2	113

1 INTRODUÇÃO

A ação de confiar permite-nos delegar tarefas e interagir, permitindo-nos alcançar determinados objetivos que dependem de ações colaborativas para serem realizados. Assim sendo, a confiança é alvo de estudos em diversas áreas, como psicologia, sociologia, filosofia e economia. Ostrom (1998), definiu a confiança como a expectativa de uma pessoa em relação às ações de outras pessoas que afetam suas escolhas.

Segundo López et al. (2011), que propôs um modelo unificador do conceito de confiança para o campo das Ciências Humanas, a confiança é a expectativa de um indivíduo ou organização de que os compromissos firmados com outros indivíduos ou organizações possam ser cumpridos. Tal modelo foi proposto levando-se em consideração as características de diversos modelos presentes na literatura.

Em Ciências da Computação, a área de Inteligência Artificial Distribuída (IAD), atua na resolução de problemas distribuídos, dividindo um problema em problemas menores, a fim de tornar a solução menos complexa (FERBER, 1999). Dentro da IAD, temos a sub-área de Sistemas Multiagente (SMA), que são sistemas compostos por entidades computacionais, conhecidas como agentes, os quais interagem entre si, para realizar as tarefas que lhes foram atribuídas (WOOLDRIDGE, 2009); (FERBER, 1999). Assim como no campo das Ciências Humanas, o estudo da confiança também está presente na área da IAD, em especial nos SMA (LU et al., 2009); (SABATER; SIERRA, 2005); (SABATER-MIR; PAOLUCCI, 2007) e (SENSOY et al., 2013).

Os SMA podem ser abertos - permitindo a entrada e/ou saída dessas entidades no ambiente a qualquer momento - ou fechados - que não permitem a entrada e/ou saída de agentes após a criação do sistema. Tanto nos ambientes abertos, como nos fechados, as interações entre os agentes, normalmente, ocorrem através da delegação de tarefas, que podem ser cumpridas de maneira satisfatória ou não pelos agentes. Através dessa capacidade de interação, o sistema consegue evoluir e resolver os problemas para os quais ele foi projetado.

Para permitir que um SMA tenha o comportamento esperado, os agentes (ou o sistema) precisam ter a capacidade de mensurar a qualidade das informações trocadas entre si, para considerar somente aquelas que conduzirão o sistema para um estado de sucesso. Assim, as ideias de confiança e reputação entre agentes tornam-se aspectos importantes na interação e cooperação dos SMA: sua utilização visa mitigar a troca de informações indesejáveis ou inconsistentes entre os agentes.

A definição de confiança e reputação entre os autores não é clara

(TEACY et al., 2006); (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004) e (SABATER; SIERRA, 2005). São utilizadas diferentes nomenclaturas para representar a confiança e reputação. Lu et al. (2009) sugere que a confiança pode ser entendida como a capacidade de um agente confiar em outro delegando parte de suas tarefas e a reputação como um conjunto de opiniões formadas por um grupo de agentes a respeito de um determinado agente. Na Seção 3.5.1 é dada uma definição de confiança e reputação que será utilizada no escopo desse trabalho.

Devido a capacidade dos SMA abertos receberem novos agentes a qualquer momento, é difícil para um agente verificar se um agente recém ingressado no sistema é confiável, uma vez que, existe pouca informação referente ao novo agente. Para tentar amenizar essa dificuldade, vários modelos de confiança e reputação vem sendo desenvolvidos na área de SMA (SABATER; SIERRA, 2001); (TEACY et al., 2006); (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004); (SABATER; SIERRA, 2005) e (GUEMKAM et al., 2013). Levando-se em consideração que os sistemas computacionais, cada vez mais, estão tornando-se sistemas de larga escala, abertos, dinâmicos e distribuídos, contendo uma grande quantidade de agentes que agem por interesse próprio (TEACY et al., 2006), a utilização da confiança e da reputação nesses sistemas torna-se aspecto fundamental para que haja uma efetiva interação entre os agentes (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004).

Através de métricas de confiança, é possível tornar os resultados obtidos através das interações entre os agentes mais seguros e confiáveis. Devido às características de flexibilidade e adaptabilidade dos SMA, essa abordagem de Inteligência Artificial Distribuída, vem sendo bastante utilizada em sistemas distribuídos abertos - lojas virtuais, bolsa de valores, Wikipédia, aplicações distribuídas específicas - que precisam obter soluções para seus problemas de forma eficaz e com um bom grau de certeza (SABATER; SIERRA, 2005).

Assim como a confiança, que até pouco tempo atrás era estudada apenas pelas áreas das Ciências Humanas (Psicologia, Economia, entre outras), a emoção também vem sendo estudada pelos cientistas da computação. Tal estudo visa a criação de sistemas computacionais que possuam uma maior interação emocional com os usuários ou consigo mesmo, como no caso dos SMA, em que o aspecto emocional ocorre entre os próprios agentes (ADAM et al., 2006); (DASTANI; MEYER,); (MEYER, 2006); (STEUNEBRINK; DASTANI; MEYER, 2007a); (LORINI, 2008) e (TURRINI; MEYER; CASTELFRANCHI, 2007).

A emoção é um dos aspectos mais centrais e penetrantes da experiência humana, possuindo várias facetas, como sentimentos e experiência;

fisiologia e comportamento; cognições e conceituações, estando relacionada diretamente com ações, eventos e objetos (ORTONY; CLORE; COLLINS, 1990). Devido a essas características, acredita-se que a emoção possa ser incorporada à agentes, tornando o seu comportamento mais próximo ao comportamento humano, a fim de imitar alguns aspectos do ser humano na forma como este toma certas decisões, como por exemplo, na hora de confiar.

1.1 FORMULAÇÃO DO PROBLEMA

Existem vários modelos de confiança entre agentes na literatura, entre eles, Regret (SABATER; SIERRA, 2005); Travos (TEACY et al., 2006); FIRE (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004) e Marsh (MARSH, 1994). Esses modelos, de um modo geral, adotam uma abordagem quantitativa, não levando em conta o caráter simbólico que caracteriza a cognição humana. Informações de natureza subjetiva que são influenciadas pelas emoções, e outras circunstâncias, são desconsideradas. Portanto, esses modelos não levam em consideração a subjetividade presente na maneira como nós humanos confiamos ou não uns nos outros (LU et al., 2009). Além disso, utilizam valores arbitrários em alguns momentos do cálculo da confiança, tornando parte da ação de confiar independente do comportamento do agente e portanto dissociada do contexto no qual o agente está inserido.

Uma vez que as emoções estão fortemente relacionadas com o ato de confiar e que a Inteligência Artificial busca aproximar o comportamento da máquina ao comportamento humano, acredita-se que a proposta de um modelo de confiança que incorpora emoções, que considere, não somente dados quantitativos, como também informações simbólicas que possam lidar com percepções subjetivas da realidade e de modo especial com determinadas emoções vivenciadas pelo agente, possa trazer bons resultados, comparado aos modelos existentes na literatura. O modelo proposto tem como diferencial uma maior subjetividade do agente no ato de confiar, aproximando o seu modo de tomar decisões, em relação a confiança nos demais agentes, mais similar aos humanos.

O modelo proposto gerencia as avaliações e as emoções dos agentes que estão relacionadas ao processo do cálculo confiança, porém os agentes podem sentir emoções provenientes de fontes externas, que não são pertinentes ao modelo. Por exemplo, uma arquitetura EBDI pode fornecer diferentes tipos de emoções ao modelo proposto, permitindo que o cálculo de confiança leve em consideração tanto as emoções internas ao modelo quanto as emoções externas.

Uma vez que o modelo está em um nível mais abstrato do as arqui-

teturas de agentes, ele tem como propósito realizar o cálculo da confiança e gerenciar as emoções relacionadas a esse processo. A forma como ele pode comunicar-se com uma arquitetura de agentes e como os agentes se comunicam entre si não estão no escopo desse trabalho.

1.2 OBJETIVO

Esse trabalho tem como objetivo propor um modelo de confiança entre agentes baseado em emoções, que utiliza informações de natureza simbólica, que estão relacionadas diretamente com o contexto no qual o agente está inserido.

1.2.1 Objetivos Específicos

- Estudar o estado da arte dos modelos de confiança na tomada de decisões a partir da confiança mútua entre agentes.
- Estudar o modelo de emoções OCC aplicado à área de Humanas.
- Estudar os modelos de emoções em agentes.
- Formalizar as emoções e das intensidades das emoções utilizadas no modelo.
- Propor um modelo de confiança entre agentes utilizando emoções.
- Avaliar o modelo proposto.

1.3 METODOLOGIA

A realização desse trabalho foi conduzida de maneira experimental, através de etapas que possibilitaram atingir os objetivos definidos. Primeiramente foi feito um levantamento do estado da arte onde foram estudados e definidos alguns conceitos principais sobre confiança, reputação e emoções em agentes. Também foram estudados e analisados alguns modelos de confiança/reputação e de emoção para agentes. Os critérios considerados na escolha dos modelos estudados foram:

- Relevância na área;
- Abrangência;

- Escolher modelos que utilizem diferentes abordagens para calcular a confiança em agentes;

Para obter o referencial teórico foi realizada uma pesquisa nas seguintes bases de dados:

- <http://scholar.google.com.br>
- <http://www.scopus.com/home.url>
- <http://ieeexplore.ieee.org/Xplore/home.jsp>

Os termos pesquisados mais relevantes foram: “*trust agents*”, “*reputation agents*”, “*emotions agents*” e “*trust reputation models*”.

Realizado o levantamento bibliográfico, foi definido um modelo que agrega emoções a um modelo de confiança para agentes. Para a escolha do modelo de confiança foram levadas em consideração a abrangência do modelo, seu funcionamento e a possibilidade de agregar emoções a sua estrutura. Entre os modelos de emoções analisados, foi escolhido um que represente de forma quantitativa as emoções para que essas possam ser inseridas no modelo de confiança.

Por fim, o modelo foi avaliado através de simulações de dois cenários distintos. Nessas simulações foram analisados diferentes conjuntos de dados que demonstram a viabilidade computacional do modelo proposto nesse trabalho.

1.4 ORGANIZAÇÃO DO TRABALHO

Esse trabalho está organizada da seguinte forma: o próximo Capítulo descreve conceitos e características de agentes e SMA; no Capítulo 3 são apresentados modelos de confiança em SMA; no Capítulo 4 é apresentado o conceito de emoções em agentes e alguns trabalhos na área; o Capítulo 5 apresenta e formaliza o modelo proposto nessa dissertação e apresenta uma validação do modelo. Por fim, o Capítulo 6 conclui o trabalho.

2 AGENTES E SMA

Neste capítulo são apresentados os principais conceitos na área de Agentes e Sistemas Multiagente que norteiam este trabalho.

2.1 AGENTES

Um agente é uma entidade computacional, dirigida por um conjunto de objetivos próprios, que está situada em um ambiente específico, no qual atua de forma flexível e autônoma a fim de atingir seus objetivos (WOOLDRIDGE, 2009); (FERBER, 1999); (FOSTER; KESSELMAN, 2004).

Segundo (WOOLDRIDGE, 2009), os agentes podem exibir três tipos de comportamento:

- Reativo: Mantém uma interação constante com o ambiente, e responde às alterações que ocorrem nele a fim de satisfazer seus objetivos, não tendo a capacidade de evoluir e aprender com novas interações.
- Pró-ativo (Deliberativos): Apresentam um comportamento orientado a objetivos, tomando a iniciativa para satisfazê-los. Portanto, diferentemente dos agentes reativos, não são conduzidos somente por eventos do ambiente, mas também por iniciativa própria.
- Social: Possuem habilidade para interagir com outros agentes através da cooperação, coordenação e negociação, ou seja, são capazes de se comunicar com outros. Essa habilidade é fundamental nos SMA, uma vez que, alguns objetivos serão alcançados, de forma eficaz, somente quando houver interação entre vários agentes.

Uma das mais importantes arquiteturas para a concepção de agentes deliberativos é a arquitetura BDI (acrônimo de *Belief, Desire and Intention* - Crença, Desejo e Intenção). Essa arquitetura reflete o processo de raciocínio prático dos humanos e utiliza estados mentais para modelar o comportamento do agente. Devido a essas características essa arquitetura é amplamente utilizada em modelos de emoções em agentes (LORINI, 2008); (JIANG; VIDAL; HUHS, 2007); (ADAM et al., 2006) e (STEUNEBRINK; DASTANI; MEYER, 2007a).

2.2 SISTEMAS MULTIAGENTE

Os SMA são sistemas compostos por diversos agentes, inseridos em um ambiente, que interagem entre si para solucionar um determinado problema (WOOLDRIDGE, 2009). Esses sistemas têm sido uma boa alternativa na resolução de problemas em ambientes abertos e distribuídos, devido a sua flexibilidade e capacidade de interação entre suas entidades (agentes).

2.2.1 Propriedades dos Ambientes MA

Devido a grande diversidade de problemas tratados pela IAD, há uma vasta variabilidade de ambientes MA, entretanto, pode-se classificar esses ambientes em categorias bem definidas. A escolha do ambiente correto é fundamental para que o sistema comporte-se da maneira desejada (RUSSELL; NORVIG, 2004). Os ambientes podem ser classificados como:

- Completamente observável *versus* Parcialmente observável: Um ambiente é dito ser completamente observável quando o agente pode obter de forma precisa, completa e atualizada as informações sobre o estado do ambiente. A maioria dos ambientes moderadamente complexos são parcialmente observáveis. Quanto mais observável for o ambiente, mais simples será a tarefa de projetar os agentes que atuarão nele, uma vez que, cada agente terá acesso a um volume maior de informações do sistema, tornando suas ações mais simples de serem tomadas, já que o agente terá conhecimento de grande parte das variáveis presentes no ambiente.
- Determinístico *versus* Não Determinístico: Em um ambiente determinístico cada ação possui um único efeito, ou seja, para cada ação realizada, sabe-se, com certeza, para qual estado o ambiente irá. Já nos ambientes não determinísticos para cada ação executada há mais de um possível estado resultante, não tendo-se certeza absoluta das modificações que irão ocorrer no ambiente após a realização de um ação. Logo, essa incerteza torna o desenvolvimento de agentes para esse tipo de ambiente mais complexo.
- Estático *versus* Dinâmico : Assume-se que um ambiente estático permanecerá inalterado, exceto pelas ações executadas pelos agentes. Já nos ambientes dinâmicos, existem outros processos atuando sobre o

ambiente, portanto mudanças podem ocorrer não somente pelas ações dos agentes.

Outra classificação importante dos ambientes trata da mobilidade dos agentes. Os ambientes podem ser classificados em abertos ou fechados. Nos ambientes fechados todos os agentes são conhecidos no momento da criação do ambiente, não sendo permitida a entrada de novos agentes após a inicialização do sistema. Já nos ambientes abertos, os agentes podem entrar e sair a qualquer momento, e cada agente é inserido no ambiente por diferentes partes interessadas, que possuem diferentes metas e objetivos (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004). Assume-se, portanto, que em um SMA aberto:

1. Os agentes tendem a ser autônomos, porém pouco confiáveis;
2. Nenhum agente pode saber tudo sobre o ambiente, ou seja, o ambiente é parcialmente observável;
3. Os agentes não são controlados por nenhuma unidade central.

Considerando as características acima citadas, o tipo mais complexo de ambiente, é aquele que é parcialmente observável, não determinístico, dinâmico e aberto.

O modelo proposto nesse trabalho busca ser o mais flexível possível, a fim de abranger a maioria dos ambientes existentes, principalmente os que são mais complexos, uma vez que os sistemas computacionais estão tornando-se cada vez mais abertos, dinâmicos e distribuídos (TEACY et al., 2006).

3 MODELOS DE CONFIANÇA EM SMA

Nesse capítulo, inicialmente serão citadas algumas vantagens que a confiança e a reputação agregam aos SMA, em seguida será feita uma revisão literária e uma análise dos modelos de confiança mais relevantes da literatura. Ao final, serão feitas algumas considerações sobre os modelos estudados, e introduzida uma definição clara de reputação e confiança, assumida neste trabalho.

Segundo Lu et al. (2009), ao introduzir confiança e reputação em um SMA, alguns benefícios podem ser obtidos, tais como:

1. Eliminação de várias comunicações desnecessárias, melhorando significativamente o desempenho dos SMA;
2. Um agente pode tomar uma decisão mais facilmente baseada na avaliação do seu grau de confiança, logo a confiança é um grande benefício em relação à teoria de decisão tradicional;
3. A confiança é um tipo de segurança básica que complementa a segurança tradicional, como criptografia, autorização e autenticação. Para um agente ser considerado seguro e eficaz ele deve possuir esses dois tipos de segurança.

Há dois tipos de modelos existentes de um ponto de vista dimensional: modelo baseado em confiança local (dimensão individual) e modelo baseado em reputação (dimensão social). Sendo que alguns modelos misturam essas duas dimensões. No modelo Spora (ZACHARIA; MOUKAS; MAES, 2000), considera-se somente a reputação dos agentes, não levando em conta a confiança local de cada agente. Modelos como Regret (SABATER; SIERRA, 2001), Referral Network (YU; SINGH, 2003) e Travos (TEACY et al., 2006) levam em consideração tanto a confiança local como a reputação, tais modelos combinam os valores da confiança local e da reputação com pesos relativos para obter o grau de confiança. Alguns modelos, como o FIRE (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004), introduzem dois conceitos adicionais: confiança baseada em papéis e reputação certificada.

A seguir são citados quatro modelos de confiança em agentes. Primeiramente, é citado o modelo de Marsh (MARSH, 1994), que é um dos primeiros modelos desenvolvidos sobre confiança entre agentes e um dos primeiros a definir o conceito de confiança direta entre agentes. A seguir, é apresentado o modelo Travos (TEACY et al., 2006) que é um modelo que usa uma abordagem probabilística que difere da abordagem dos demais modelos

estudados. A seguir, é apresentado o modelo FIRE (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004), e por último, é citado o modelo Regret (SABATER; SIERRA, 2001), que é o modelo no qual é baseado o modelo proposto.

Os modelos apresentados a seguir, utilizam nomenclaturas diferentes para representar a confiança e a reputação. Para facilitar a comparação entre esses modelos, a tabela 1 apresenta uma equivalência entre essas nomenclaturas.

Tabela 1 – Nomenclaturas

	<i>Confiança</i>	<i>Reputação</i>
Marsh	Confiança local	Nao há
Regret	Reputação Individual	Reputação Social
Fire	Confiança por Interação	Reputação por Testemunho
Travos	Confiança	Reputação

Ao final do capítulo é feita uma análise sobre esses quatro modelos.

3.1 MODELO DE MARSH

O modelo proposto por (MARSH, 1994) foi um dos primeiros modelos desenvolvidos sobre confiança local (considera apenas a interação direta entre agentes para medir a confiança). Sua arquitetura é distribuída, uma vez que, cada agente é o responsável pelo cálculo da sua confiança perante os outros agentes. Ele diferencia três tipos de confiança:

- **Confiança Básica:** Representa a *disposição* de um agente confiar em outro, sendo calculada a partir de todas as experiências acumuladas pelo agente. Experiências boas levam a uma maior disposição para confiar, e vice-versa. Usa-se a notação T_x^t para representar a disposição do agente x confiar em outro agente qualquer no tempo t .
- **Confiança Geral:** A confiança que um agente tem em outro sem levar em conta qualquer situação específica. É denotada por $T_x(y)^t$ representando a confiança geral que um agente x tem em um agente y no tempo t .
- **Confiança Situacional:** Valor da confiança que um agente tem em relação a outro, levando em consideração uma situação específica. A *utilidade* da situação, sua *importância* e a “Confiança Geral” são

os elementos considerados no cálculo da “Confiança Situacional”. A fórmula básica usada para calcular essa confiança é:

$$T_x(y, \alpha)^t = U_x(\alpha)^t \times I_x(\alpha)^t \times \widehat{T_x(y)}^t \quad (3.1)$$

onde x é o agente avaliador, y o agente sendo avaliado e α a situação. $U_x(\alpha)^t$ representa a utilidade que x obtém da situação α , $I_x(\alpha)^t$ é a importância da situação α para o agente x e $\widehat{T_x(y)}^t$ é a estimativa da confiança geral após serem levados em consideração todos os possíveis dados relevantes a respeito de $T_x(y, \alpha)$ no passado, ou seja, se t é o tempo atual, x irá agregar todas as situações $T_x(y, \sigma)^T$ com $\theta < T < t$ e σ similar ou idêntica a situação atual α . As variáveis θ e t definem a janela temporal que o agente está considerando. Portanto somente as experiências dentro dessa janela serão consideradas na agregação.

Para definir $\widehat{T_x(y)}$ o autor propôs três métodos estatísticos: média, máximo e mínimo. Cada método é identificado com um tipo diferente de agente:

- Otimista: assume o maior (máximo) valor de confiança dentre todos os valores de suas experiências passadas.
- Pessimista: assume o menor (mínimo) valor de confiança.
- Realista: calcula o valor como uma média usando a fórmula $\widehat{T_x(y)} = \frac{1}{|A|} \sum_{\alpha \in A} T_x(y, \alpha)$ onde A é o conjunto de situações similares a situação atual α avaliada na janela temporal.

O cálculo do risco e da competência é obtido a partir dos três tipos de confiança: básica, geral e situacional. O valor de confiança é utilizado para ajudar um agente a decidir se deve ou não cooperar com outro agente. Além da confiança em si, o mecanismo de decisão desse modelo leva em consideração a importância da ação a ser realizada, o risco associado à situação e a percepção da competência do agente sendo avaliado. Uma característica importante desse modelo é a de reciprocidade, ou seja, se um agente x ajudou um agente y no passado, e y nega-se a ajudar o agente x posteriormente, a confiança entre ambos irá diminuir (SABATER; SIERRA, 2005).

3.2 MODELO TRAVOS

No modelo de confiança e reputação Travos (TEACY et al., 2006) um agente a_{tr} (*truster*) possui dois métodos para calcular a confiança em outro

agente a_{te} (*trustee*) em um contexto específico. Primeiramente o agente a_{tr} faz a avaliação baseado nas interações diretas com o agente a_{te} , depois o agente a_{tr} avalia a confiabilidade de a_{te} através da reputação de a_{te} .

Nesse modelo os autores denotam o conjunto de todos os agentes como $A = \{a_1, a_2, \dots, a_n\}$, sendo n o número total de agentes do SMA. Com o passar do tempo, pares distintos de agentes $\{a_x, a_y\} \subseteq A$ podem interagir entre si, dirigidos pelos contratos que especificam as obrigações de cada agente com seu parceiro de interação. Além disso, assume-se, que todas as interações ocorrem sob obrigações semelhantes, já que, um agente pode comportar-se de maneira diferente quando solicitado a fornecer um tipo diferente de serviço. Então a melhor forma de indicar como um agente irá se comportar em suas interações futuras é através da avaliação de obrigações similares realizadas por ele no passado. Uma interação entre um agente avaliador (*truster*) $a_{tr} \in A$ e um agente sendo avaliado (*trustee*) $a_{te} \in A$, é considerada bem sucedida por a_{tr} se a_{te} cumpriu com suas obrigações. O resultado de uma interação entre a_{tr} e a_{te} é resumida a uma variável binária $O_{a_{tr}, a_{te}}$, onde $O_{a_{tr}, a_{te}} = 1$ indica sucesso (e $O_{a_{tr}, a_{te}} = 0$ indica insucesso) do ponto de vista do agente a_{tr} . O resultado observado no tempo t é denotado por $O_{a_{tr}, a_{te}}^t$, e o conjunto de todos os resultados observados do tempo 1 até o tempo t é denotado por $O_{a_{tr}, a_{te}}^{1:t}$, com $t > 0$. Para cada valor de t , no máximo uma interação entre dois agentes quaisquer poderá ocorrer. Portanto, $O_{a_{tr}, a_{te}}^{1:t}$ é um conjunto de no máximo t variáveis binárias representando todas as interações que ocorreram entre a_{tr} e a_{te} até o tempo t (incluindo t).

$$O_{a_{tr}, a_{te}}^{1:t} = \begin{cases} 1 & \text{se o contrato é cumprido por } a_{tr} \\ 0 & \text{caso contrário} \end{cases} \quad (3.2)$$

A qualquer momento do tempo t , o histórico das interações entre a_{tr} e a_{te} é registrada como uma tupla $R_{a_{tr}, a_{te}}^t = (m_{a_{tr}, a_{te}}^t, n_{a_{tr}, a_{te}}^t)$ onde o valor de $m_{a_{tr}, a_{te}}^t$ é o número de interações bem sucedidas do agente a_{tr} com a_{te} , enquanto $n_{a_{tr}, a_{te}}^t$ é o número de interações mal sucedidas. A tendência de um agente a_{te} cumprir ou não com suas obrigações é dirigida pelo seu comportamento, que é representado pela variável $B_{a_{tr}, a_{te}} \in [0, 1]$. Assim, $B_{a_{tr}, a_{te}}$ representa a probabilidade que a_{te} irá cumprir com suas obrigações durante a interação com a_{tr} . Por exemplo, se $B_{a_{tr}, a_{te}} = 0.8$ é esperado que em 80% das vezes, o agente a_{te} irá cumprir seus contratos com o agente a_{tr} , resultando em um total de 80% de interações bem sucedidas entre ambos.

$$B_{a_{tr}, a_{te}} = p(O_{a_{tr}, a_{tr}} = 1), \text{ tal que } B_{a_{tr}, a_{te}} \in [0, 1] \quad (3.3)$$

A confiança de um agente a_{tr} em um agente a_{te} , denotado por $\tau_{a_{tr}, a_{te}}$, é a avaliação de a_{tr} da probabilidade de que a_{te} irá cumprir com suas obrigações

durante uma interação. A confiabilidade *confidence* de a_{tr} nessa avaliação sobre a_{te} é denotada por $\gamma_{a_{tr},a_{te}}$, representando a precisão do valor da confiança calculado a partir do número de observações (*evidências*) usadas no cálculo da confiança. Assim, quanto maior o número de *evidências*, maior será a confiabilidade da informação.

3.2.1 Confiança e Confiabilidade

Nesse modelo, os autores adotam uma abordagem probabilística para modelar a confiança, baseada nas experiências passadas de um agente sendo avaliado. Se um agente avaliador (a_{tr}), tem todas as informações sobre o agente sendo avaliado (a_{te}), então, de acordo com a_{tr} , a probabilidade de a_{te} cumprir com suas obrigações é expressa por $B_{a_{tr},a_{te}}$. Contudo, normalmente não se tem toda a informação necessária sobre a_{te} , logo, foi utilizado o *valor esperado* de $B_{a_{tr},a_{te}}$ dado o conhecimento (conjunto de todos os resultados das interações observadas) de a_{tr} . Assim, define-se o nível de confiança $\tau_{a_{tr},a_{te}}$ no tempo t como o *valor esperado* de $B_{a_{tr},a_{te}}$ dado o conjunto dos resultados $O_{a_{tr},a_{te}}^{1:t}$, através da fórmula:

$$\tau_{a_{tr},a_{te}} = E[B_{a_{tr},a_{te}} | O_{a_{tr},a_{te}}^{1:t}] \quad (3.4)$$

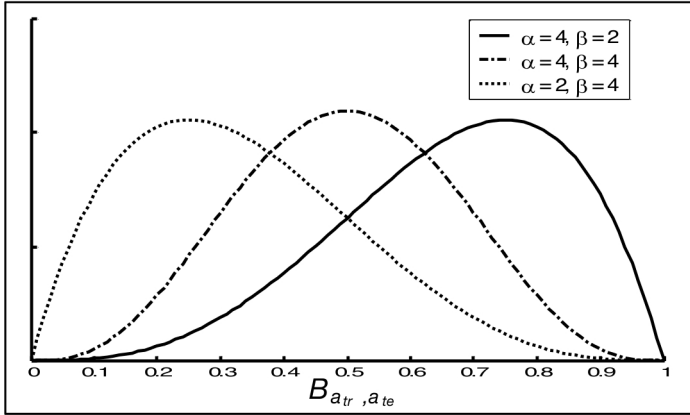
Para determinar esse *valor esperado*, utiliza-se uma distribuição probabilística definida por uma função de densidade de probabilidade (pdf), que é usada para modelar a probabilidade relativa de modo que $B_{a_{tr},a_{te}}$ tenha um valor determinado. Uma vez que a família beta das pdfs são utilizadas em distribuições para variáveis aleatórias que estão no intervalo $[0, 1]$, e por serem utilizadas em outros trabalhos sobre confiança, os autores utilizaram essa família de funções nesse modelo (TEACY et al., 2006).

A fórmula padrão para distribuições beta é dada a seguir, na qual os dois parâmetros α e β definem a forma da função densidade no gráfico (Figura 1). Nesse gráfico, o eixo horizontal representa os possíveis valores de $B_{a_{tr},a_{te}}$, e o eixo vertical fornece a probabilidade *relativa* de que cada um desses valores seja verdadeiro para $B_{a_{tr},a_{te}}$. O valor mais provável para $B_{a_{tr},a_{te}}$ é o valor máximo da curva, já o grau de incerteza em relação a veracidade de $B_{a_{tr},a_{te}}$ é expresso pelo formato da curva. Se os valores de α e β estiverem próximos a 1, o gráfico será mais alongado, indicando um alto grau de incerteza sobre $B_{a_{tr},a_{te}}$.

$$f(B_{a_{tr},a_{te}} | \alpha, \beta) = \frac{(B_{a_{tr},a_{te}})^{\alpha-1} (1 - B_{a_{tr},a_{te}})^{\beta-1}}{\int_0^1 U^{\alpha-1} (1 - U)^{\beta-1} dU}, \quad (3.5)$$

onde $\alpha, \beta > 0$.

Figura 1 – Exemplo de gráfico de uma função beta (TEACY et al., 2006)



Para calcular-se o valor de $\tau_{a_{tr}, a_{te}}$ baseado no resultados das interações observadas por a_{tr} , primeiramente precisa-se encontrar os valores de α e β que representam as crenças de a_{tr} . Supondo-se que, antes de observar qualquer resultado das interações de a_{tr} com a_{te} , acredita-se que todos os possíveis valores para $B_{a_{tr}}$ são possivelmente iguais, então a configuração inicial de α e β para a_{tr} é $\alpha = \beta = 1$.

$$\alpha = m_{a_{tr}, a_{te}}^{1:t} + 1 \quad \text{e} \quad \beta = n_{a_{tr}, a_{te}}^{1:t} + 1 \quad (3.6)$$

onde t é o tempo da avaliação.

Nessa equação, a configuração dos parâmetros é obtida adicionando-se o número de resultados bem sucedidos à configuração inicial de α , e o número de resultados mal sucedidos à β .

Por fim, o valor final de $\tau_{a_{tr}, a_{te}}$ é calculado aplicando-se a equação padrão (3.7) para o *valor esperado* de uma distribuição beta para essa

configuração de parâmetros.

$$E[B_{a_{tr},a_{te}}|\alpha,\beta] = \frac{\alpha}{\alpha + \beta} \quad (3.7)$$

O valor da confiança $\tau_{a_{tr},a_{te}}$ não faz distinção entre os casos em que o agente a_{tr} possui a informação adequada sobre a_{te} e os casos em que ele não possui. Por isso, utiliza-se uma medida de confiabilidade do valor de confiança (3.8), denotada por $\gamma_{a_{tr},a_{te}}$, representando a probabilidade posterior de que o valor atual de $B_{a_{tr},a_{te}}$ está dentro de uma margem de erro aceitável ε sobre $\tau_{a_{tr},a_{te}}$.

$$\gamma_{a_{tr},a_{te}} = \frac{\int_{\tau_{a_{tr},a_{te}} - \varepsilon}^{\tau_{a_{tr},a_{te}} + \varepsilon} X^{\alpha-1} (1-X)^{\beta-1} dX}{\int_0^1 U^{\alpha-1} (1-U)^{\beta-1} dU} \quad (3.8)$$

Essa equação pode ser interpretada como a proporção da distribuição de probabilidade que está entre os limites $(\tau_{a_{tr},a_{te}} - \varepsilon)$ e $(\tau_{a_{tr},a_{te}} + \varepsilon)$. O erro ε influencia no valor da confiabilidade que um agente calcula para um determinado conjunto de observações, assim, um maior valor de ε provoca um decréscimo de grande parte da distribuição beta no intervalo $[\tau_{a_{tr},a_{te}} - \varepsilon, \tau_{a_{tr},a_{te}} + \varepsilon]$, resultando em um maior valor para $\gamma_{a_{tr},a_{te}}$.

3.2.2 Reputação

O modelo Travos trata a reputação de forma semelhante aos outros modelos analisados (SABATER; SIERRA, 2001) e (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004), sendo necessária sua utilização quando um agente a_{tr} quer fazer uma avaliação sobre o comportamento de um agente a_{te} , e possui poucas informações sobre ele ou a confiança direta entre ambos possui uma baixa confiabilidade. Assim, a reputação pode aumentar a precisão da confiabilidade do valor de confiança que a_{tr} tem em a_{te} , auxiliando a_{tr} na sua tomada de decisão.

Os valores de confiabilidade são utilizados nesse processo de tomada de decisão, levando o agente a buscar mais evidências quando necessário. O agente a_{tr} calcula o valor da confiança ($\tau_{a_{tr},a_{te}}$) baseado em suas experiências diretas com a_{te} . Caso o valor da confiança tenha uma confiabilidade abaixo de um valor mínimo pré-determinado, denotado por θ_γ , então o agente a_{tr} irá solicitar as opiniões dos outros agentes sobre a_{te} , a fim de tornar o valor de sua confiabilidade maior que θ_γ . A reputação de a_{te} é formada por essas opiniões provenientes de outros agentes, e através delas o agente a_{tr} pode obter um grande conjunto de observações sobre a_{te} .

Os autores definem uma opinião *verdadeira* de uma fonte $a_{op} \in A$ no tempo t , sobre o agente a_{te} , como uma tupla $R_{a_{op},a_{te}}^t = (m_{a_{op},a_{te}}^t, n_{a_{op},a_{te}}^t)$. Por sua vez, a opinião *relatada* por uma fonte a_{op} sobre a_{te} é denotada por $\hat{R}_{a_{op},a_{te}}^t = (\hat{m}_{a_{op},a_{te}}^t, \hat{n}_{a_{op},a_{te}}^t)$. Essa distinção torna-se importante porque o agente a_{op} pode informar de forma não verdadeira sua opinião $\hat{R}_{a_{op},a_{te}}^t$.

Um agente a_{tr} precisa formular um único valor de confiança a partir de todas as opiniões recebidas. Assim, assumindo-se que cada opinião é independente, uma forma de unir essas opiniões é enumerar as interações bem sucedidas e mal sucedidas de todos os relatórios recebidos, onde p é o número total de relatos (equação (3.9)). Os valores resultantes, denotados por $N_{a_{tr},a_{te}}$ e $M_{a_{tr},a_{te}}$, representam a reputação de a_{te} do ponto de vista do agente a_{tr} . Esses valores são usados para calcular os parâmetros (equação (3.10)) para uma distribuição beta, resultando em um valor de confiança determinado pelas opiniões dos outros agentes. Além disso, o agente a_{tr} considera cada experiência direta tida com a_{te} adicionando seus próprios valores para $n_{a_{tr},a_{te}}$ e $m_{a_{tr},a_{te}}$ na equação.

$$N_{a_{tr},a_{te}} = \sum_{k=0}^p \hat{n}_{a_k,a_{te}}, \quad M_{a_{tr},a_{te}} = \sum_{k=0}^p \hat{m}_{a_k,a_{te}} \quad (3.9)$$

$$\alpha = M_{a_{tr},a_{te}} + 1 \quad \text{e} \quad \beta = N_{a_{tr},a_{te}} + 1 \quad (3.10)$$

Asseguradas as seguintes condições:

- **Condição 1 (Comportamento comum):** O comportamento do agente a_{te} deve ser independente da identidade do agente a_{tr} com o qual ele está interagindo. Assim:

$$\forall a_{te} \quad \forall a_{op}, \quad B_{a_{tr},a_{te}} = B_{a_{tr},a_{op}}$$

- **Condição 2 (Dizer a verdade):** O fornecedor da reputação deve relatar suas observações com exatidão e verdade. Então:

$$\forall a_{te} \quad \forall a_{op}, \quad R_{a_{op},a_{te}}^t = \hat{R}_{a_{op},a_{te}}^t$$

deseja-se que o valor resultante da confiança e da confiabilidade sejam semelhantes, mesmo que todas as observações tivessem sido feitas diretamente pelo próprio agente a_{tr} . Contudo, isso também implica o modo como diferentes agentes avaliam o comportamento do agente a_{te} , fazendo com que, um fornecedor de opinião para a_{tr} classifique uma interação como bem sucedida, ou mal sucedida, da mesma forma que o próprio a_{tr} .

3.2.3 Reputação Imprecisa

No modelo Travos os autores fornecem um método para lidar com relatos de reputação imprecisos, que surgem quando a Condição 1 ou 2 (acima) é quebrada, devido a opinião de um agente sendo maléfico ou um agente a_{te} agindo de forma inconsistente para diferentes agentes. Em ambos os casos, o agente deve ser capaz de avaliar a confiabilidade dos relatos passados a ele e ajustar ou ignorar as opiniões consideradas pouco confiáveis. Para filtrar essas opiniões imprecisas primeiramente estima-se a probabilidade da precisão da opinião reletada, para posteriormente ajustar os valores das opiniões que formam a reputação.

A estimativa da probabilidade depende do valor da opinião corrente sendo considerada, denotada por $\hat{R}_{a_{op},a_{te}} = (\hat{m}_{a_{op},a_{te}}, \hat{n}_{a_{op},a_{te}})$. Se E^r é o valor esperado de uma distribuição beta D^r , de tal forma que $\alpha^r = \hat{m}_{a_{op},a_{te}} + 1$ e $\beta^r = \hat{n}_{a_{op},a_{te}} + 1$, pode-se estimar a probabilidade de E^r estar dentro de uma margem de erro em torno de $B_{a_{tr},a_{te}}$, chamada de precisão de a_{op} de acordo com a_{tr} , denotado por $\rho_{a_{tr},a_{op}}$. Essa estimativa, leva em consideração os resultados de todas as interações prévias para os quais a_{op} forneceu uma opinião similar a $\hat{R}_{a_{op},a_{te}}$ sobre a_{te} para a_{tr} , para cada a_{te} . A partir desses resultados, constrói-se uma distribuição beta D^o a qual, se o valor esperado E^o é próximo a E^r , então as opiniões de a_{op} são correlacionadas com o que é realmente observado, e a precisão de a_{op} será elevada. Por outro lado, se E^r diverge significativamente de E^o , então a_{op} possui uma baixa precisão.

O processo para alcançar essa estimativa consiste em dividir o intervalo de possíveis valores de E^r em N intervalos disjuntos bin_1, \dots, bin_n , então calcula-se E^r , e encontra-se o intervalo bin^o que contém o valor E^r . Então, sendo $H_{a_{tr},a_{op}}$ o conjunto de todos os pares da forma $(O_{a_{tr},a_x}, \hat{R}_{a_{op},a_x})$, onde $a_x \in A$, e O_{a_{tr},a_x} é o resultado da interação para a qual, antes de ser observada por a_{tr} , a_{op} deu a opinião \hat{R}_{a_{op},a_x} , pode-se encontrar um subconjunto $H_{a_{tr},a_{op}}^r \subseteq H_{a_{tr},a_{op}}$, que compreende todos os pares para os quais o valor esperado da opinião esteja em bin^o . Depois, conta-se o número total de pares em $H_{a_{tr},a_{op}}^r$ para o qual o resultado da interação foi bem sucedido ($C_{success}$) e aqueles que não foram (C_{fail}), baseado nesses valores, os parâmetros para D^o são definidos como $\alpha^o = C_{success} + 1$ e $\beta^o = C_{fail} + 1$. Usando-se D^o , calcula-se $\rho_{a_{tr},a_{op}}$ como a porção da magnitude total de D^o que está no intervalo bin^o (equação (3.11)).

$$\rho_{a_{tr},a_{te}} = \frac{\int_{min(bin^o)}^{max(bin^o)} X^{\alpha^o-1} (1-X)^{\beta^o-1} dX}{\int_0^1 U^{\alpha^o-1} (1-U)^{\beta^o-1} dU} \quad (3.11)$$

Nesse modelo, cada agente executa essa operação para determinar a probabilidade da precisão das opiniões relatadas, no entanto, a quantia e o tamanho dos intervalos (*bins*) determinam uma margem de erro aceitável para a precisão do fornecedor da opinião. Assim, para intervalos (*bins*) grandes, a estimativa da precisão de um grande conjunto de fornecedores de opinião converge para 1.

Para descrever como ocorre o ajuste das opiniões, os autores (TEACY et al., 2006) introduzem algumas novas notações:

1. Seja D^c a distribuição beta resultante da combinação de todas as informações de reputação do agente a_{tr} , usando-se as equações (3.9) e (3.10).
2. Seja D^{c-r} a distribuição construída a partir das mesmas equações (3.9) e (3.10), exceto pelo fato da opinião em consideração, $\hat{R}_{a_{op}, a_{te}}$, ser omitida.
3. Seja \bar{D} o resultado do ajuste da opinião da distribuição D^r , de acordo com o processo descrito a seguir.
4. Por fim, o desvio padrão (denotado por σ), o valor esperado e os parâmetros de cada distribuição são referenciados usando seu respectivo sobrescrito, como por exemplo, D^c possui parâmetros α^c e β^c , com desvio padrão σ^c e valor esperado E^c .

Para reduzir o efeito das opiniões não confiáveis em D^c , adiciona-se $\hat{R}_{a_{op}, a_{te}}$ a uma reputação do agente a_{te} , e move-se E^c em direção a E^r . O desvio padrão de D^r contribui com o valor de confiabilidade para o valor da reputação combinado, e o seu valor relativo a σ^{c-r} determina o quão distante E^c ficará de E^r .

Os autores adotam uma abordagem que reduz significativamente os valores dos parâmetros que sejam muito altos, a menos que a probabilidade da opinião do avaliador esteja muito próximo a 1. Especificamente, reduz-se a distância entre o valor esperado e o desvio padrão de D^r , e a distância entre o valor esperado e o desvio padrão da distribuição uniforme, $\alpha = \beta = 1$, que representa um estado sem informação (equações (3.12) e (3.13)). Nessas equações denota-se o desvio padrão da distribuição uniforme como $\sigma_{uniform}$ e seu valor esperado como $E_{uniform}$. Ajustando-se o desvio padrão dessa maneira, ao invés de alterar os parâmetros α e β diretamente, assegura-se que parâmetros com valores maiores são reduzidos mais do que valores menores.

$$\bar{E} = E_{uniform} + \rho_{a_{tr}, a_{op}} \cdot (E^r - E_{uniform}) \quad (3.12)$$

$$\bar{\sigma} = \sigma_{uniform} + \rho_{a_{tr}, a_{op}} \cdot (\sigma^r - \sigma_{uniform}) \quad (3.13)$$

Depois de calcular os valores \bar{E} e $\bar{\sigma}$, utiliza-se as equações (3.14) e (3.15) para encontrar os parâmetros $\bar{\alpha}$ e $\bar{\beta}$ da distribuição ajustada, e a partir desses parâmetros calcula-se os valores ajustados para $\hat{m}_{a_{op}, a_{te}}$ e $\hat{n}_{a_{op}, a_{te}}$ denotados por $\bar{m}_{a_{op}, a_{te}}$ e $\bar{n}_{a_{op}, a_{te}}$ respectivamente (3.16). Esses dois últimos valores são utilizados no cálculo do valor da confiança combinada como na equação (3.9), e a medida que a precisão de a_{op} segundo a_{tr} , expressa por $\rho_{a_{tr}, a_{op}}$, aproxima-se de 0, $\bar{m}_{a_{op}, a_{te}}$ e $\bar{n}_{a_{op}, a_{te}}$ também irão aproximar-se de 0. Assim, se o valor de $\rho_{a_{tr}, a_{op}}$ é 0, nenhuma opinião fornecida por a_{op} irá afetar a tomada de decisão do agente a_{tr} .

$$\bar{\alpha} = \frac{\bar{E}^2 - \bar{E}^3}{\bar{\sigma}^2} - \bar{E} \quad (3.14)$$

$$\bar{\beta} = \frac{(1 - \bar{E})^2 - (1 - \bar{E})^3}{\bar{\sigma}^2} - (1 - \bar{E}) \quad (3.15)$$

$$\bar{m}_{a_{op}, a_{te}} = \bar{\alpha} - 1 \quad , \quad \bar{n}_{a_{op}, a_{te}} = \bar{\beta} - 1 \quad (3.16)$$

3.3 FIRE

O modelo FIRE, proposto por (DONG-HUYNHA; JENNINGS; SHAD-BOLT, 2004), é um modelo de confiança e reputação integrado, com uma arquitetura de tomada de decisão distribuída entre os agentes. Ele incorpora quatro fontes de informação - confiança por interação, reputação baseada em testemunho, confiança baseada em papéis e reputação certificada – as quais são combinadas para fornecer uma métrica de confiança em praticamente todas as circunstâncias. Essa variedade de fontes torna-se importante, visto que, em várias situações nem todas estarão prontamente disponíveis, além de permitir aos agentes combiná-las para lidar com as incertezas do ambiente.

3.3.1 Confiança por interação (IT)

A confiança por interação (*Interaction Trust*) é a confiança resultante da interação direta entre dois agentes, sendo equivalente à confiança direta proposta no modelo RegreT (SABATER; SIERRA, 2001). Tomando como exemplo uma compra eletrônica em que um agente a compra um produto de um agente b , o preço ou a qualidade desse produto são fatores importantes

que podem ser avaliados pelo agente a , de acordo com o serviço prestado por b . Essas avaliações são tuplas $r = (a, b, i, c, v)$, onde a e b são os agentes participantes na interação i , e v é a avaliação dada ao agente b por a para o termo c (preço ou qualidade do produto). O intervalo de v é $[-1, +1]$, onde -1 significa completamente negativo, $+1$ completamente positivo e 0 neutro ou incerto.

Para calcular a IT a partir de experiências passadas, o agente precisa gravar suas avaliações passadas em uma base de dados local. Quando o agente a está calculando o valor da IT de um agente b em relação ao termo c , ele necessita procurar no banco de dados por todas as avaliações do tipo $(a, b, _, c, _)$, onde $_$ representa qualquer valor. Esse conjunto de avaliações é representado por $R(a, b, c)$, então a IT (denotada por τ_I) é calculada como uma média ponderada dos valores de todas as avaliações no conjunto:

$$\tau_I(a, b, c) = \sum_{r_i \in R(a, b, c)} \omega(r_i) \cdot v_i \quad (3.17)$$

onde v_i é o valor da avaliação r_i e $\omega(r_i)$ é o peso correspondente a r_i . O peso $\omega(r_i)$ para cada avaliação é selecionado de forma a dar uma maior relevância às avaliações mais recentes, com a restrição que $\sum_{r_i \in R(a, b, c)} \omega(r_i) = 1$, para assegurar que o valor da confiança $\tau_I(a, b, c)$ esteja no intervalo $[-1, +1]$.

Nesse modelo, cada valor de confiança possui uma taxa de confiabilidade que reflete a certeza do modelo em produzir os valores de confiança a partir dos dados levados em consideração. Esse valor é construído a partir de duas medidas:

- $\rho_N(a, b, c)$: a medida da confiabilidade baseada no número de avaliações levadas em consideração no cálculo de τ_I . À medida que o número dessas avaliações (n) aumenta, o grau de confiabilidade aumenta até ele atingir um limite definido (denotado por m).

$$\rho_N(a, b, c) = \begin{cases} \frac{n}{m} & \text{quando } n \leq m \\ 1 & \text{quando } n > m \end{cases} \quad (3.18)$$

onde n é a cardinalidade de $R(a, b, c)$. O valor da função $\frac{n}{m}$ varia entre 0 e 1 para n no intervalo $[0, m]$. Assim, a confiabilidade $\rho_N(a, b, c)$ aumenta de 0 para 1 quando o número de avaliações n aumenta de 0 para m , e permanece em 1 quando n excede m .

- $\rho_D(a, b, c)$: a confiabilidade do desvio padrão. Quanto maior a variabilidade nos valores das avaliações, maior será a instabilidade do outro agente no cumprimento de seus acordos.

$$\rho_D(a, b, c) = 1 - \sum_{r_i \in R(a, b, c)} \frac{\omega(r_i) \cdot |v_i - \tau_I(a, b, c)|}{2} \quad (3.19)$$

Assim, a medida de confiabilidade da IT, denotado por $\rho_{\tau_I}(a, b, c)$, é definida por:

$$\rho_{\tau_I}(a, b, c) = \rho_N(a, b, c) \cdot \rho_D(a, b, c) \quad (3.20)$$

3.3.2 Confiança baseada em papéis (RT)

A confiança baseada em papéis (*Role-based Trust*) resulta das relações baseadas em papéis entre dois agentes (exemplo, entre agentes da mesma organização, ou relações derivadas de relações da vida real entre os criadores dos agentes, como amizade ou parentesco). Uma vez que não há métodos gerais para computacionalmente quantificar a confiança baseada nessas relações, os autores utilizaram regras para atribuir os valores da RT. (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004). Isso permite aos usuários finais, adicionar novas regras para personalizar a aplicação de acordo com suas necessidades. As regras são tuplas $rul = (role_a, role_b, c, v_D, e_D)$ que descrevem regras onde $role_a$ e $role_b$ são os papéis dos agentes a e b respectivamente. Então o desempenho esperado de b em uma interação com a é v_D ($v_D \in [-1, 1]$) com respeito ao termo c ; $e_D \in [0, 1]$ é o nível padrão de influência dessa regra no resultado do valor da RT.

Um exemplo de regra poderia ser:

$$rul_1 = (comprador, vendedor, qualidade, -0.5, 0.3)$$

representando a crença de um agente que um vendedor usualmente vende um produto de baixa qualidade, entretanto a confiabilidade dessa crença é baixa (0.3).

Cada agente possui seu próprio conjunto de regras que é armazenado em uma base de regras local. Para determinar a RT em relação a um agente b o agente a procura em sua base de regras, por regras que são relevantes. O valor da RT é dado por:

$$\tau_R(a, b, c) = \frac{\sum_{rul_i \in Rules(a, b, c)} e_{D_i} \cdot v_{D_i}}{\sum_{rul_i \in Rules(a, b, c)} e_{D_i}} \quad (3.21)$$

onde $rul_i = (role_a, role_b, c, v_{D_i}, e_{D_i})$ é uma regra do conjunto de regras $Rules(a, b, c)$, que é um subconjunto da base de regras, que contém somente as regras que são relevantes aos papéis de a , aos papéis de b , e ao termo c .

Uma vez que as regras para RT são especificadas pelo dono do agente, a confiabilidade da RT também é configurada por ele. O valor da confiabilidade da RT é denotado por $\rho \tau_R(a, b, c)$, variando no intervalo $[0, 1]$.

3.3.3 Reputação por testemunho (WR)

Esse tipo de reputação é obtido através da observação do comportamento de um agente b pelos outros agentes (testemunhas) do sistema (*Witness Reputation*). Quando um agente a quer avaliar a reputação de b , ele verifica com os outros agentes como foram suas experiências com b . Para que isso ocorra, o agente a precisa encontrar as testemunhas que já interagiram com o agente b . Nesse sistema, cada agente mantém uma lista de agentes conhecidos. Assim, quando se olha para uma determinada informação, um agente pode enviar a seus conhecidos uma consulta, para que esses tentem responder, caso não consigam, eles vão lhe enviar referências que apontam para agentes que eles acreditam serem capazes de possuir a informação desejada. Para auxiliar nessa busca por informação, cada agente possui uma medida do grau de probabilidade que representa a sua capacidade de responder a uma consulta, essa medida precisa ser definida de uma forma específica para cada aplicação.

A fórmula geral para WR é:

$$\tau_W(a, b, c) = \sum_{r_i \in R_W(a, b, c)} \omega(r_i) \cdot v_i \quad (3.22)$$

onde $R_W(a, b, c)$ é o conjunto de avaliações das testemunhas encontradas pelo agente a , o peso $\omega(r_i)$ para cada avaliação é definido da mesma forma que na Seção 3.3.1, e v_i é o valor da avaliação de r_i . A medida de confiabilidade para WR, denotada por $\rho \tau_W(a, b, c)$, é definida a partir das avaliações em $R_W(a, b, c)$, de maneira similar a explicada na Seção 3.3.1.

3.3.4 Reputação Certificada (CR)

A reputação certificada (*Certified Reputation*) são classificações apresentadas pelo próprio agente b sendo avaliado, que foram obtidas a partir de interações passadas com os outros agentes. Essas classificações são certificações a respeito da conduta desse agente em momentos passados, e lhe permitem provar seu possível desempenho ao agente a (avaliador). Uma vez que o agente pode escolher quais avaliações irá apresentar, um agente racional irá fornecer somente suas melhores avaliações. Portanto, a informação

proveniente da CR possivelmente irá superestimar o comportamento esperado de um agente. Então, embora não se possa garantir o desempenho do agente b em interações futuras, a CR revela uma perspectiva parcial do comportamento passado desse agente (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004). O principal benefício desse tipo de informação é a sua alta disponibilidade, portanto a CR está disponível aos agentes na maioria das vezes, mesmo nas situações em que os outros componentes (IT, RT e WR) possam falhar ao fornecer uma medida de confiança segura.

O processo da CR acontece da seguinte forma:

- Depois de cada transação, b pede aos seus parceiros para fornecerem suas avaliações sobre seu desempenho que é armazenada em sua base de dados.
- Quando a contacta b para expressar seu interesse em utilizar seus serviços, ele solicita que b forneça suas referências sobre seu desempenho passado.
- O agente a recebe as avaliações sobre b do agente b . Ele acessa a confiabilidade das avaliações e calcula o valor de confiança para b . O valor da CR $\tau_C(a, b, c)$ e sua confiabilidade $\rho\tau_C(a, b, c)$, são calculados da mesma maneira que os componentes da WR, com a diferença que a entrada é o conjunto das avaliações fornecidas pelo próprio agente b , ao invés das avaliações fornecidas pelas testemunhas.

3.3.5 Confiança Final

Os valores da confiança e reputação são combinados, formando uma equação composta que representa o provável desempenho do agente. Tanto no cálculo do valor da confiança $\tau(a, b, c)$ quanto no da confiabilidade $\rho\tau(a, b, c)$ utiliza-se uma média ponderada entre as quatro fontes de informação.

$$\tau(a, b, c) = \frac{\sum_{k \in \{I, R, W, C\}} \omega_k \cdot \tau_k(a, b, c)}{\sum_{k \in \{I, R, W, C\}} \omega_k} \quad (3.23)$$

$$\rho\tau(a, b, c) = \frac{\sum_{k \in \{I, R, W, C\}} \omega_k}{\sum_{k \in \{I, R, W, C\}} W_k} \quad (3.24)$$

onde $\omega_k = W_k \cdot \rho\tau_k(a, b, c)$, e W_I, W_R, W_W, W_C são os coeficientes correspondentes aos componentes IT, RT, WR e CR respectivamente. Esses coeficientes são estabelecidos pelo usuário final para refletir a importância de cada um

desses componentes na sua aplicação.

3.4 REGRET

O modelo de confiança e reputação Regret, proposto por (SABATER; SIERRA, 2001), é um sistema modular voltado para pequenos e médios ambientes de comércio eletrônico onde a relação social entre indivíduos é importante.

A reputação é vista como uma opinião ou visão de um agente sobre algo. Cada agente possui uma opinião diferente sobre os demais, e portanto, a reputação está, de certa maneira, ligada à subjetividade. Essa opinião é formada e atualizada ao longo do tempo através das interações com os outros agentes do sistema. Essas interações fornecem como resultado *impressões* que são registradas pelos agentes. As *impressões* são a base que permite valorar a reputação dos agentes.

Os autores avaliam a confiança de um forma totalmente descentralizada, onde cada agente é capaz de mensurar por si só, a reputação dos demais. Levando-se em consideração que cada agente pertence a um determinado grupo, são definidas três dimensões de reputação: individual, social e ontológica:

- Individual: Representa a situação em que o agente só leva em consideração suas experiências próprias para calcular a confiança. Pode ser entendida como a confiança local do modelo de Marsh.
- Social: Representa às experiências prévias dos membros de um determinado grupo em relação a um agente específico. Equivale ao conceito da palavra reputação conforme entendemos na nossa língua.
- Ontológica: Essa dimensão considera a reputação do agente um conceito de múltiplos aspectos, sendo resultante da combinação de diferentes “pedaços” de informações, e não simplesmente um conceito simples e abstrato. O modo como esses “pedaços” são combinados para construir a reputação é chamado de dimensão ontológica. Por exemplo, a reputação de uma boa companhia aérea engloba a reputação de ter bons aviões, a reputação de nunca extraviar a bagagem e a reputação de servir boas refeições. Por sua vez, a reputação de ter bons aviões engloba a reputação de ter um bom serviço de manutenção e a reputação de frequentemente renovar a frota. É importante salientar que cada indivíduo possui sua própria estrutura ontológica (Ex. alguns passageiros podem considerar que a qualidade das refeições não

influência na reputação de uma companhia aérea), permitindo que a combinação e a maneira de avaliar a importância da reputação seja feita de forma individual.

Nas equações pertencentes a esse modelo, os grupos dos agentes são denotados por letras maiúsculas (A, B, \dots), e os agentes por letras minúsculas indexadas (a_1, b_2, \dots). Por exemplo, um agente denotado por a_1 pertence ao grupo A . E, por fim, o conjunto de todos os agentes é denotado pela letra \mathbf{A} .

3.4.1 Resultado e Impressões

O resultado (*outcome*) de um diálogo entre dois agentes é um contrato inicial que estabelece os termos e condições de uma transação, juntamente com o resultado das ações tomadas ou os valores atuais dos termos da transação. O conjunto de todos possíveis *resultados* é denotado pela letra \mathbf{O} . Por exemplo, um *resultado* de um diálogo relacionado a uma transação comercial, do ponto de vista do agente comprador b , poderia ser:

$$o_b = (DataEntrega_c = 10/02 \wedge Preco_c = 2000 \wedge Qualidade_c = A \wedge DataEntrega = 15/02 \wedge Preco = 2000 \wedge Qualidade = C)$$

já o mesmo diálogo, do ponto de vista do agente vendedor a , poderia gerar o seguinte *resultado*:

$$o_a = (DataEntrega_c = 10/02 \wedge Preco_c = 2000 \wedge DataEntrega = 15/02 \wedge Preco = 2000)$$

Nesse exemplo, as variáveis com o subscrito c representam o acordo inicial entre ambos os agentes. Assim o agente b esperava um produto com qualidade A (boa) porém recebeu um produto de qualidade C (ruim), além de ter recebido o produto com 5 dias de atraso.

Uma impressão é uma avaliação subjetiva feita por um agente sobre um certo aspecto de um *resultado*. Uma impressão ι é representada por uma tupla na forma:

$$\iota = (a, b, o, \varphi, t, W) \quad (3.25)$$

onde $a, b \in \mathbf{A}$ são os agentes que dialogam (sendo a o agente que está julgando), $o \in \mathbf{O}$ é o *outcome*, φ a variável do *resultado* que está sendo julgada, t é o tempo em que a *impressão* foi registrada e $W \in [-1, 1]$ representa a opinião subjetiva do agente a com respeito a φ .

O conjunto de todas as possíveis impressões é denotado por I e a base de dados das impressões do agente a é denotada por $IDB^a \subseteq I$. Define-se $IDB_p^a \subseteq IDB^a$ como o conjunto das impressões em IDB^a que satisfazem o padrão p , onde a forma geral de um padrão é:

$$(a, b, o, \phi, t, W) | \text{condicao} \quad (3.26)$$

com a *condicao* sendo uma fórmula lógica de primeira ordem sobre os componentes da impressão.

3.4.2 Reputação Individual

A Reputação Individual (RI) representa a confiança que um determinado agente tem em outro, levando em consideração apenas as suas impressões, sem considerar a opinião dos outros agentes.

O valor de RI é calculado diretamente da base de impressões do agente através da equação (3.27) que utiliza uma média ponderada atribuindo às impressões mais recentes um peso maior.

$$R^i(IDB_p^a) = \sum_{i_i \in IDB_p^a} \rho(t, t_i) \cdot W_i \quad (3.27)$$

onde $p = \{(a, b, _, \phi, _) | \text{true}\}$, “ $_$ ” denota um valor qualquer, t é o tempo corrente e IDB_p^a é o conjunto das impressões do agente a que satisfazem o padrão p . Essa equação realiza um somatório dos produtos entre os valores das impressões (W_i) e os pesos $\rho(t, t_i)$. A notação $R_{a \rightarrow b}(\phi)$ é utilizada para representar $R^i(IDB_p^a)$.

O peso $\rho(t, t_i)$ é definido através da função $\frac{f(t_i, t)}{\sum_{i_j \in IDB_p^a} f(t_j, t)}$. Essa função possui valores normalizados e atribui valores maiores a valores mais próximos de t , ou seja, quando mais recente for a avaliação dada pelo agente a , maior será o valor de $\rho(t, t_i)$. Um exemplo de função $f(t_i, t)$ pode ser $f(t_i, t) = \frac{t_i}{t}$. Uma vez que $W_i \in [-1, 1]$ e $\rho(t, t_i)$ são valores normalizados, o valor de $R_{a \rightarrow b}(\phi) \in [-1, 1]$.

Para exemplificar, vamos supor que a quer calcular a RI em relação a b e que $t = 2$. Sendo $IDB_p^a = \{0.5_{t=1}, -0.2_{t=2}\}$, temos que $\sum_{i_j \in IDB_p^a} f(t_j, t) = \frac{1}{2} + \frac{2}{2} = 1.5$. Logo $\rho(2, 1) = \frac{f(1, 2)}{1.5} = \frac{0.5}{1.5} = 0.33$ e $\rho(2, 2) = \frac{f(2, 2)}{1.5} = 0.67$. Então o valor de $R_{a \rightarrow b}(\phi)$ é obtido através do somatório $(0.33) \cdot (0.5) + (0.67) \cdot (-0.2) = 0.165 - 0.134 = 0.031$.

Para cada valor de Reputação Individual, uma medida de confiabilidade é associada, a qual caracteriza a relevância desses valores. Essa medida

é obtida a partir de duas métricas: o número de *impressões* usadas no processo do cálculo dos valores de reputação $Ni(IDB_p^a)$ (3.28), e a variabilidade dos valores dessas *impressões* expressa por $Dt(IDB_p^a)$ (3.29).

$$Ni(IDB_p^a) = \begin{cases} \sin(\frac{\pi}{2.itm}|IDB_p^a|) & \text{se } |IDB_p^a| \in [0, itm] \\ 1 & \text{caso contrário} \end{cases} \quad (3.28)$$

onde $|IDB_p^a|$ é a cardinalidade de IDB_p^a , ou seja, o número de impressões usadas para calcular a reputação, e itm é o limite máximo do grau de confiabilidade, tendo um valor específico para cada aplicação.

$$Dt(IDB_p^a) = 1 - \sum_{t_i \in IDB_p^a} \rho(t, t_i) \cdot |W_i - R^t(IDB_p^a)| \quad (3.29)$$

com $Dt(IDB_p^a) \in [0, 1]$, onde valores próximos a 0 indicam uma baixa credibilidade do valor da reputação, enquanto valores próximos a 1 indicam uma alta credibilidade no valor da reputação.

Por fim, a confiabilidade de uma reputação individual, denotada por $RL_{a \rightarrow b}(\varphi)$, é expressa pela fórmula:

$$RL(IDB_p^a) = (1 - \mu) \cdot Ni(IDB_p^a) + \mu \cdot Dt(IDB_p^a) \quad (3.30)$$

com $\mu \in [0, 1]$.

3.4.3 Reputação Social

Diferentemente da reputação individual, a Reputação Social (RS) leva em consideração três fontes de informação:

1. A interação do agente a (avaliador) com os membros do grupo que o agente b (avaliado) pertence;
2. O que os membros do grupo A (o grupo do agente a) pensam sobre o agente b ;
3. O que os membros do grupo A pensam sobre o outro grupo B (o grupo do agente b).

Essa dimensão social dá ao agente a possibilidade de modelar uma característica importante das sociedades complexas: a relação entre os grupos (SABATER; SIERRA, 2001).

A reputação oriunda da interação do agente a com os outros membros do grupo B é dada pela equação:

$$R_{a \rightarrow B}(\varphi) = \sum_{b_i \in B} \omega^{ab_i} . R_{a \rightarrow b_i}(\varphi) \quad (3.31)$$

onde $\sum_{b_i \in B} \omega^{ab_i} = 1$.

Os pesos ω^{ab_i} representam a importância da avaliação dada pelo agente a em relação a cada um dos agentes b . Quando o agente a sabe o quão próximo (do ponto de vista comportamental) o agente b é dos outros membros do grupo, o agente a pode ajustar o valor de ω^{ab_i} dando um peso maior para os agentes próximos a b . Quando a incerteza sobre essa proximidade é máxima $\omega^{ab_i} = \frac{1}{|B|}$.

Da mesma forma que na reputação individual, a confiabilidade dessa reputação é dada por:

$$RL_{a \rightarrow B}(\varphi) = \sum_{b_i \in B} \omega^{ab_i} . RL_{a \rightarrow b_i}(\varphi) \quad (3.32)$$

onde $RL_{a \rightarrow b_i}$ representa os valores de confiabilidade da reputação individual (ver equação 3.30) do agente a em relação a cada um dos agentes do grupo B . O valor de $RL_{a \rightarrow B}(\varphi)$ representa a relevância do valor de $R_{a \rightarrow B}(\varphi)$.

Para representar a opinião dos membros do grupo A sobre o agente b e a confiabilidade dessa reputação, utiliza-se as fórmulas:

$$R_{A \rightarrow b}(\varphi) = \sum_{a_i \in A} \omega^{a_i b} . R_{a_i \rightarrow b}(\varphi) \quad (3.33)$$

$$RL_{A \rightarrow b}(\varphi) = \sum_{a_i \in A} \omega^{a_i b} . RL_{a_i \rightarrow b}(\varphi) \quad (3.34)$$

onde $\sum_{a_i \in A} \omega^{a_i b} = 1$.

Por fim, para expressar o que os membros do grupo A pensam sobre o outro grupo B , utiliza-se as fórmulas:

$$R_{A \rightarrow B}(\varphi) = \sum_{a_i \in A} \omega^{a_i B} . R_{a_i \rightarrow B}(\varphi) \quad (3.35)$$

$$RL_{A \rightarrow B}(\varphi) = \sum_{a_i \in A} \omega^{a_i B} . RL_{a_i \rightarrow B}(\varphi) \quad (3.36)$$

onde $\sum_{a_i \in A} \omega^{a_i B} = 1$.

Os valores $\omega^{a_i b}$ e $\omega^{a_i B}$ podem refletir a credibilidade ou relações hierárquicas dentro do grupo, dando uma maior importância às informações vindas dos agentes com maior credibilidade ou com uma maior posição na hierarquia.

O valor final da confiança é calculado combinando-se o valor da RI (3.27) com os três valores da RS (3.31), (3.33) e (3.35).

$$SR_{a \rightarrow b}(\varphi) = \xi_{ab} \cdot R_{a \rightarrow b}(\varphi) + \xi_{aB} \cdot R_{a \rightarrow B}(\varphi) + \xi_{Ab} \cdot R_{A \rightarrow b}(\varphi) + \xi_{AB} \cdot R_{A \rightarrow B}(\varphi) \quad (3.37)$$

O valor de $SR_{a \rightarrow b}(\varphi)$ representa a confiança de a em relação a b . Através da análise desse valor, a irá decidir se confia ou não em b .

A confiabilidade $SRL_{a \rightarrow b}$, que caracteriza a relevância do valor de $SR_{a \rightarrow b}(\varphi)$, ou seja, o quão confiável é o valor de $SR_{a \rightarrow b}(\varphi)$, é obtida através da equação:

$$SRL_{a \rightarrow b}(\varphi) = \xi_{ab} \cdot RL_{a \rightarrow b}(\varphi) + \xi_{aB} \cdot RL_{a \rightarrow B}(\varphi) + \xi_{Ab} \cdot RL_{A \rightarrow b}(\varphi) + \xi_{AB} \cdot RL_{A \rightarrow B}(\varphi) \quad (3.38)$$

onde $\xi_{ab} + \xi_{aB} + \xi_{Ab} + \xi_{AB} = 1$.

Ajustando esses parâmetros, o agente pode dar uma maior ou menor importância a cada fonte de opinião. Esses valores não são necessariamente estáticos e podem mudar ao longo do tempo de acordo com o estado mental do agente. Por exemplo, um agente que foi, recentemente, prejudicado por ter confiado na opinião de outros agentes, pode dar um peso maior ao parâmetro ξ_{ab} nos próximos cálculos de reputação.

3.4.4 Dimensão Ontológica

Nesse modelo, os autores consideram que a confiança e a reputação não estão ligadas a um único aspecto. Através da dimensão ontológica, é possível combinar os valores das reputações de diferentes aspectos para calcular valores mais complexos. Cada agente pode ter uma estrutura ontológica distinta, representada como um grafo cíclico, que combina a reputação individual e social de maneiras diferentes, para avaliar a importância desses valores quando eles são combinados. Para calcular-se a reputação de um nodo no grafo, agrega-se a reputação de cada filho que, por sua vez, é recursivamente obtido a partir de seus filhos (SABATER; SIERRA, 2001). A reputação de um nodo i em um grafo ontológico é calculada pela fórmula:

$$OR_{a \rightarrow b}(i) = \begin{cases} \sum_{j \in \text{children}(i)} \omega_{ij} \cdot OR_{a \rightarrow b}(j) & \text{se } \text{children}(i) \neq \emptyset \\ SR_{a \rightarrow b}(j) & \text{caso contrário} \end{cases} \quad (3.39)$$

E sua confiabilidade por:

$$ORL_{a \rightarrow b}(i) = \begin{cases} \sum_{j \in children(i)} \omega_{ij} \cdot ORL_{a \rightarrow b}(j) & \text{se } children(i) \neq \emptyset \\ SRL_{a \rightarrow b}(j) & \text{caso contrário} \end{cases} \quad (3.40)$$

3.5 CONSIDERAÇÕES SOBRE OS MODELOS

Nessa Seção serão feitas algumas considerações sobre os modelos estudados, e posteriormente será introduzida uma definição sobre confiança e reputação adotada neste trabalho.

O modelo de Marsh (MARSH, 1994) foi um dos primeiros modelos desenvolvidos sobre confiança local. Ele leva em consideração apenas a interação direta entre agentes para medir a confiança (confiança direta), ou seja, ele não faz uso da reputação, por esse fato, ele é limitado em relação a capacidade de calcular a confiança quando os agentes nunca interagiram entre si. Sua apresentação nesse trabalho é importante devido ao seu pioneirismo na área, e a sua definição de confiança direta entre agentes.

O modelo de confiança Travos (TEACY et al., 2006) utiliza uma abordagem probabilística para calcular a confiança. Foi desenvolvido para ser utilizado, principalmente, em SMA abertos e assume que o comportamento dos agentes não mudam com o tempo, no entanto, essa hipótese nem sempre é verdadeira. A principal característica - que também está presente em outros modelos estudados Regret (SABATER; SIERRA, 2001) e FIRE (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004) - desse modelo é a possibilidade de um agente avaliar a confiança em outro agente de forma direta, através de experiências passadas, ou através da reputação, quando tais experiências não existirem ou forem insuficientes. Ao estabelecer a confiança em outros agentes e escolher aquele que lhe é mais confiável, o agente tem a capacidade de maximizar a probabilidade de que sua interação seja bem sucedida. Outra característica presente é a possibilidade de filtrar as opiniões imprecisas de outros agentes, permitindo que a reputação seja utilizada para aumentar significativamente o desempenho do sistema, já que as informações indesejáveis serão descartadas.

Assim como o modelo Regret, o modelo FIRE (DONG-HUYNHA; JENNINGS; SHADBOLT, 2004), é um modelo de reputação com uma arquitetura de tomada de decisão distribuída entre os agentes. Além das duas fontes de informação presentes no Regret - RI (Reputação Individual) e RS (Reputação Social) - que no modelo FIRE são chamadas respectivamente de - IT (Confiança por Interação) e WR (Reputação por Testemunho), ele

incorpora duas novas fontes de informação (Confiança baseada em Papéis e Reputação Certificada). Essas quatro fontes são combinadas para fornecer uma métrica de confiança em praticamente todas as circunstâncias.

O modelo Regret (SABATER; SIERRA, 2001) possui uma abordagem descentralizada que permite aos agentes avaliar o grau de confiança sem precisar recorrer a um mecanismo centralizado e modela, de certa maneira, a subjetividade presente na ação de um agente confiar ou não em outro. Devido a natureza modular desse modelo, o agente pode decidir se utiliza a Reputação Individual (RI) juntamente com a Reputação Social (RS) ou apenas uma delas, para tomar a decisão de confiar ou não. Como o modelo leva em consideração várias fontes de informações no cálculo da confiança, ele possui uma elevada capacidade de adaptação, mesmo com diferentes graus de conhecimento entre os agentes, permitindo assim, que o sistema seja utilizado em vários cenários.

Entre os modelos apresentados, os que mais se adequam a proposta dessa dissertação, em um primeiro momento, são os modelos Regret e FIRE. Eles possibilitam que alguns pesos utilizados em suas funções sejam substituídos por símbolos emocionais, representados pelas intensidades das emoções vivenciadas pelos agentes.

Devido a similaridade entre esses dois modelos, ambos poderiam ter sido escolhidos. Contudo optou-se pelo Regret em um primeiro momento pela sua abrangência, simplicidade e clareza. O modelo Fire poderia também ser adotado para incorporar emoções, bastando para tal, realizar alguns ajustes para considerar as intensidades das emoções em seus cálculos de confiança e reputação.

3.5.1 Confiança versus Reputação

Buscando clarificar e definir os termos de confiança e reputação, usados nos modelos estudados, serão utilizadas definições encontradas no dicionário Michaelis (MICHAELIS..., 2012). Nesse dicionário o termo *confiança* é definido como: "Ação de confiar. Segurança íntima com que se procede". Por sua vez o termo *confiar* é definido como: "Ter confiança, acreditar. Incumbir, encarregar." O termo *reputação* é definido como: "Ato ou efeito de reputar. Fama, renome. Conceito em que uma pessoa é tida; bom ou mau nome". E o termo *reputar* é definido como: "Dar reputação ou bom nome a. Avaliar, estimar".

Levando em consideração essas definições, podemos dizer que a *reputação* é uma avaliação da conduta de uma determinada pessoa perante outras, e que *confiança* é o ato de acreditar em outra pessoa, de incumbir

uma pessoa com uma determinada tarefa. Nós podemos confiar em alguém, mesmo sem saber sobre sua reputação, portanto a reputação "vem antes da confiança", ela é um conceito que nos auxilia na decisão de confiar ou não em alguém, principalmente quando desconhecemos essa pessoa.

Analisando as considerações acima, podemos definir *confiança local*, *reputação* e *confiança* em SMA da seguinte forma:

- **Confiança Local** - É baseada na relação direta entre dois agentes e é obtida através de experiências e interações passadas entre ambos. Quando um agente *a* considerar somente a confiança local, a ação dele confiar ou não em um agente *b* dependerá somente da análise de suas interações, não sendo levada em consideração a reputação de *b*.
- **Reputação** - Ato de avaliar a conduta/fama de um determinado agente perante as opiniões de outros agentes. Portanto a reputação permite que *a* confie em *b*, sem a necessidade de ambos terem tido alguma experiência e/ou interação no passado.
- **Confiança** - É o ato final de confiar ou não em outro agente, de incumbir outro agente de uma tarefa. Esse ato de confiar pode ser baseado tanto na confiança local quanto na reputação, ou em ambas em conjunto.

Apesar de haver uma diferença de nomenclaturas entre os modelos, a idéia principal dos modelos que utilizam a reputação, é permitir ao agente confiar ou não em outro, de forma consistente, mesmo quando não houve interações passadas entre ambos.

4 EMOÇÕES EM AGENTES INTELIGENTES

O estudo das emoções está presente em várias áreas de conhecimento como Psicologia, Economia, Neurociência Cognitiva e, nos últimos anos, esse estudo também está presente nas pesquisas em IA e Ciência da Computação. Tal estudo visa a criação de sistemas de interação emocional, como por exemplo, robôs com comportamento emocional e agentes virtuais para entreterimento (LORINI, 2008) e, hoje em dia, mais e mais agentes artificiais integram habilidades emocionais para obterem expressividade, adaptabilidade, credibilidade (ADAM et al., 2006). Em estudos psicológicos, as emoções que influenciam a deliberação e o raciocínio prático de um agente são consideradas como heurísticas para prevenir uma deliberação excessiva (STEUNEBRINK; DASTANI; MEYER, 2007a).

As emoções começaram a receber uma maior atenção nas pesquisas de IA, principalmente aquelas que buscam adicionar sentimentos às interações entre humanos e máquinas. Trabalhos recentes estão investigando o raciocínio do senso comum, mas a pesquisa em aplicação de emoções para a tomada de decisão em agentes ainda é muito limitada. Essa limitação ocorre devido à complexidade de modelar o comportamento emocional do ser humano; o desconhecimento sobre o que dispara as emoções nos seres humanos e como cada um vê e sente de forma diferente uma mesma situação. Essas dificuldades também estão presentes na modelagem de emoções em agentes artificiais (JIANG; VIDAL; HUHNS, 2007). Para contornar essa limitação, alguns pesquisadores da área de SMA têm trabalhado no desenvolvimento de estruturas lógicas para especificação formal de emoções (ADAM et al., 2006); (DASTANI; MEYER,); (MEYER, 2006); (STEUNEBRINK; DASTANI; MEYER, 2007a); (LORINI, 2008) e (TURRINI; MEYER; CASTELFRANCHI, 2007).

A seguir será apresentado o modelo de emoções OCC (ORTONY; CLORE; COLLINS, 1990) e posteriormente serão apresentados alguns trabalhos que modelam as emoções em agentes.

Devido a utilização das notações (eventos, ações e objetos) que também são usadas em modelos de agentes, o modelo OCC é popular entre pesquisadores que estão construindo sistemas de raciocínio sobre emoções ou que incorporam emoções em agentes artificiais (STEUNEBRINK; DASTANI; MEYER, 2009); (ADAM et al., 2006).

4.1 MODELO OCC

O modelo psicológico de emoções OCC, proposto por (ORTONY; CLORE; COLLINS, 1990), classifica 22 tipos de emoções, positivas e negativas (figura 2). Nele, os aspectos quantitativos das emoções são descritos em termos de potencialidades, limiares e intensidades. Para cada uma das 22 emoções, o modelo fornece uma lista de variáveis que afetam a intensidade dessa emoção e se a condição que a provoca é assegurada.

Figura 2 – Especificações dos tipos de emoções do modelo OCC (ORTONY; CLORE; COLLINS, 1990) [tradução nossa]

Alegria: (satisfeito com) um evento desejável.
Angústia: (insatisfeito com) um evento indesejável.
Feliz-por: (satisfeito com) um evento que presume-se ser desejável para alguma pessoa.
Pena: (insatisfeito com) um evento que presume-se ser indesejável para alguma pessoa.
Regozijo: (satisfeito com) um evento que presume-se ser indesejável para alguma pessoa.
Ressentimento: (insatisfeito com) um evento que presume-se ser desejável para alguma pessoa.
Esperança: (satisfeito com) a expectativa de um evento desejável.
Medo: (insatisfeito com) a expectativa de um evento indesejável.
Satisfação: (satisfeito com) a confirmação da expectativa de um evento desejável.
Medo-confirmado: (insatisfeito com) a confirmação da expectativa de um evento indesejável.
Alívio: (satisfeito com) a não confirmação da expectativa de um evento indesejável.
Desapontamento: (insatisfeito com) a não confirmação da expectativa de um evento desejável.
Orgulho: (aprovação da) sua própria ação louvável.
Vergonha: (desaprovação de) sua própria ação culpável.
Admiração: (aprovação da) ação louvável de outra pessoa.
Reprovação: (desaprovação da) ação culpável de outra pessoa.
Gratificação: (aprovação da) sua própria ação louvável e (estar satisfeito com) o evento desejável relacionado.
Remorso: (desaprovação da) sua própria ação culpável e (estar insatisfeito com) o evento indesejável relacionado.
Gratidão: (aprovação da) ação louvável de outra pessoa e (estar satisfeito com) o evento desejável relacionado.
Raiva: (desaprovação da) ação culpável de outra pessoa e (estar insatisfeito com) o evento indesejável relacionado.
Amor: (gostar) de um objeto agradável.
Ódio: (não gostar) de um objeto desagradável.

4.1.1 Tipos de emoções do modelo OCC

Conforme dito anteriormente, o modelo OCC possui 22 tipos de emoções, sendo que cada uma delas é especificada através de três elementos:

1. Especificação do TIPO - Quais condições que provocam o tipo de emoção em questão.

2. **TOKENS** - Representa quais palavras podem ser classificadas como pertencentes ao tipo de emoção em questão.
3. **Variáveis que afetam a INTENSIDADE**: Estas variáveis são locais para o tipo de emoção em questão, ou seja, as variáveis globais (Ex. excitação) que afetam todas as emoções não estão incluídas. Quanto maior os valores dessas variáveis maior será a intensidade da emoção.

Uma possível especificação da classe de emoções MEDO ("FEAR"), utilizando esses três elementos, pode ser definida como:

- Especificação do TIPO: (insatisfeito com) a perspectiva de um evento indesejável.
- **TOKENS**: assustado, aterrorizado, apreensivo.
- **Variáveis que afetam a INTENSIDADE**:
 - O grau em que o evento é indesejável.
 - A probabilidade do evento.
- **EXEMPLO**: O empregado, suspeitando que não seria mais necessário para a empresa, temia ser demitido.

Por sua vez a especificação de ALEGRIA ("JOY") pode ser descrita como:

- Especificação do TIPO: (satisfeito com) um evento desejável.
- **TOKENS**: prazer, felicidade, contentamento.
- **Variáveis que afetam a INTENSIDADE**:
 - O grau em que o evento é desejável.
- **EXEMPLO**: A criança estava contente por ter ganho uma bicicleta de natal.

Como pode ser visto, na classe ALEGRIA não existe a variável de INTENSIDADE "A probabilidade do evento", já que o sentimento de ALEGRIA é em relação a um evento que já aconteceu ou que está acontecendo, diferentemente do MEDO, que é um sentimento sobre algo que pode acontecer.

A emoção de ADMIRAÇÃO ("ADMIRATION") pode ser definida como:

- Especificação do TIPO: (aprovação de) uma ação louvável de outro agente.
- TOKENS: consideração, respeito, contemplação.
- Variáveis que afetam a INTENSIDADE:
 - O grau em que a ação é louvável.
 - A força da identificação com o causador da ação.
 - A expectativa em relação a ação ser merecedora de admiração ou não.
- EXEMPLO: O pai admirou o ato heróico do filho.

Analisando essas três emoções, nota-se que as duas primeiras estão relacionadas a ocorrência ou não de um determinado evento, enquanto a última está relacionada a ação realizada por um determinado agente. Essa, e outras diferenças, serão explicadas a seguir com a apresentação da hierarquia do modelo OCC. Cada uma das 22 emoções é subdividida em subgrupos de acordo com os estímulos que a provoca.

4.1.2 Hierarquia de emoções do OCC

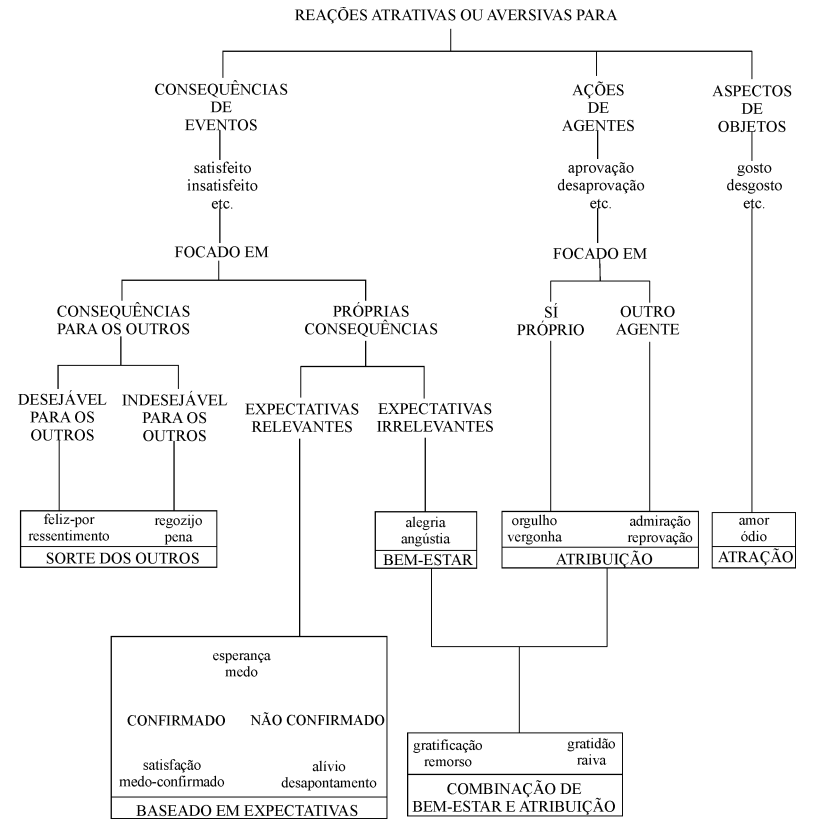
A hierarquia do modelo OCC (figura 3) possui três ramificações, cada uma representa a avaliação de um tipo diferente de estímulo em relação a determinadas atitudes e determinada avaliação de uma variável. Essas ramificações são subdivididas em grupos de emoções provocadas por condições semelhantes, e são classificadas em:

- Emoções baseadas em eventos (Ex. alegria e compaixão);
- Emoções baseadas em agentes (Ex. orgulho e vergonha);
- Emoções baseadas em aspectos de objetos (Ex. amor e ódio).

Analisando a hierarquia do modelo, nota-se que as emoções *satisfeito*, *insatisfeito*, *aprovação*, *desaprovação*, *gosto* e *desgosto* são a base para todas as outras emoções.

A seguir serão apresentadas as 3 ramificações do modelo (ORTONY; CLORE; COLLINS, 1990). Ao longo dessa dissertação, será dada uma atenção maior às emoções baseadas em eventos e baseadas em agentes, por serem essenciais na modelagem de emoções em agentes e por estarem

Figura 3 – Estrutura de emoções do modelo OCC (ORTONY; CLORE; COLLINS, 1990) [tradução nossa]



fortemente relacionadas aos eventos e às ações dos próprios agentes. Assim, as emoções baseadas em aspectos de objetos, serão apenas citadas, uma vez que não possuem uma relevância tão forte como os outros dois tipos de emoções.

4.1.3 Emoções baseadas em eventos

As emoções presentes nesse ramo são induzidas a partir da avaliação de um evento, com relação aos objetivos do agente. O *desejo*, do evento

ocorrer ou não, é a principal variável de intensidade nesse tipo de emoção, pois o agente só vai desejar que ocorram os eventos que auxiliem nos seus objetivos. Assim, se o evento ocorrido era desejado pelo agente, este sentirá emoções positivas, caso contrário, sentirá emoções negativas. Esse ramo é subdividido em:

- **Emoções de bem-estar** - É um sentimento que só depende do desejo de que um determinado evento ocorra. Exemplo: O agente sentiu alegria por ter ganho dinheiro.
- **Emoções baseadas em expectativa** - São focadas no *desejo* do próprio agente de que um evento esperado (incerto) venha a acontecer. Nesse tipo de emoções, utiliza-se uma variável de intensidade local, que representa a *probabilidade* do evento ocorrer. Exemplo: O agente tem esperança de que irá ganhar dinheiro.
- **Emoções baseadas na prosperidade dos outros** - São provocadas a partir do suposto desejo de outro agente. Nessas emoções são utilizadas três variáveis de intensidade local, *desejo do outro*, *merecimento* e *amizade*. O *desejo do outro* é a avaliação de quão esperado é o evento para o outro, o *merecimento* representa quanto o agente *a* acredita que o agente *b* mereça o que lhe ocorreu, e a *amizade* representa a afinidade entre *a* e *b*. Exemplo: O agente *a* ficou ressentido porque o agente *b* perdeu dinheiro.

4.1.4 Emoções baseadas em agentes

Essas emoções são induzidas a partir do julgamento de uma ação, em relação aos "padrões morais", ou seja, uma ação é louvável quando ela segue os padrões. As emoções desse grupo referem-se tanto a ação do próprio agente quanto a ação de um outro agente.

- **Emoções de atribuição** - São focadas na aprovação da ação de um agente, nesse tipo de ação há duas variáveis de intensidade local, *força da unidade* e *desvio da expectativa*. A primeira intervém nas próprias emoções do agente para representar o grau em que o agente se identifica com o autor da ação, e a segunda representa o grau em que a ação executada difere do que é normalmente esperado, de acordo com seu papel social. Exemplo: O agente sentiu vergonha por ter roubado dinheiro.

4.1.5 Emoções baseadas em aspectos de objetos

- **Emoções de atração** - Essas emoções estão relacionadas pela atração que um agente sente por algo ou alguém, nesse tipo de emoção há apenas uma variável de intensidade local *familiaridade*. De acordo com o modelo OCC, quanto mais familiar for um objeto agradável ao agente, mais ele irá amá-lo, por sua vez, quanto mais familiar for um objeto desagradável, maior será o ódio por ele. Exemplo: O agente *a* ama o agente *b*.

4.1.6 Emoções compostas

- **Emoções de atração** Adicionalmente, algumas dessas ramificações são combinadas para formar um grupo de emoções compostas, baseadas na consequência de eventos e ações, sendo o resultado da combinação de emoções de bem-estar e emoções de atribuição. Exemplo: O agente *a* sentiu gratidão pelo agente *b* por ele ter devolvido seu dinheiro perdido.

4.2 MODELOS DE EMOÇÕES EM AGENTES

Nesse capítulo serão apresentados alguns trabalhos relacionados a modelagem de emoções em agentes. A maioria desses trabalhos modelam as emoções de forma qualitativa, entretanto alguns autores sugerem uma abordagem quantitativa das emoções através da atribuição de valores numéricos às intensidades das emoções (STEUNEBRINK; DASTANI; MEYER, 2007b).

4.2.1 Modelo de Lorini

O modelo de Lorini (LORINI, 2008) é um modelo formal baseado na lógica de agentes BDI, para a especificação de emoções complexas, através da especificação de crenças e objetivos em que as emoções (tais como, surpresa, esperança, medo, decepção, alívio) possam ser formalizadas.

Tomando como exemplo uma lógica BDI simples com operadores para a crença, desejo e tempo linear, e dada uma fórmula arbitrária ϕ , o operador Bel_i é usado para expressar o que um determinado agente *i* acredita, o operador $Goal_i$ é usado para expressar o que um determinado agente *i* quer alcançar, o operador X é usado para expressar os fatos que serão verdadeiros

no próximo estado do sistema e $Poss_i$ é definido como a dualidade de Bel_i .

- $Bel_i\phi$ representa a crença do agente i em ϕ ;
- $Goal_i\phi$ representa que o agente i quer que ϕ seja verdade;
- $X\phi$ representa que ϕ será verdadeiro no próximo estado do sistema.
- $Poss_i\phi$ representa que o agente i pensa que ϕ é possível.
- $Poss_i\phi \Leftrightarrow \neg Bel_i\neg\phi$.

Através desses operadores é possível formalizar algumas emoções básicas, como por exemplo, alegria (*joy*) e esperança (*hope*):

$$Joy_i\phi \Leftrightarrow Bel_i\phi \wedge Goal_i\phi \quad (4.1)$$

De acordo com esta definição, o agente i sente alegria sobre ϕ se e somente se i acredita que ϕ é verdade e quer que ϕ seja verdade. Neste sentido, i fica contente com o fato dele acreditar que tenha alcançado um resultado desejável.

$$Hope_i\phi \Leftrightarrow Poss_iX\phi \wedge Goal_iX\phi \quad (4.2)$$

Essa definição representa a esperança do agente i na veracidade de ϕ no próximo estado se e somente se i achar que ϕ possivelmente será verdadeira no próximo estado e i quer que ϕ seja verdadeira no próximo estado.

Segundo o autor, ainda não existe na literatura um modelo formal que consiga de forma adequada caracterizar as emoções complexas, como arrependimento, vergonha, culpa. Essas emoções são tidas como complexas por basearem-se em formas de raciocínios complexas e em um conjunto maior de variáveis de avaliações do que o conjunto envolvido nas "emoções básicas", e por isso são difíceis de serem formalizadas.

Na literatura da Psicologia enfatiza-se que tais emoções envolvem formas muito sofisticadas de raciocínio, tais como, auto-atribuição de responsabilidade, raciocínio contrafactual, raciocínio sobre normas e ideais (LORINI, 2008). Assim sendo, uma lógica de emoções complexas precisa levar em consideração algumas importantes dimensões, são elas:

Dimensão contrafactual: Emoções complexas como arrependimento e culpa são baseados em um raciocínio contrafactual sobre suas próprias emoções e ações. Por exemplo, um agente sente-se arrependido se e somente se ele acredita não ter conseguido um de seus objetivos e acredita que, se ele decidisse fazer algo diferente, ele teria alcançado seu objetivo. Neste sentido,

existem emoções complexas que se baseiam na capacidade de imaginar cenários alternativos ao atual, os quais poderiam terem sido obtidos se algo diferente tivesse sido feito. Um agente se sente arrependido não apenas quando pensa que a sua decisão de fazer uma determinada ação levou a um resultado ruim, mas também, quando ele pensa que a sua decisão de evitar realizar uma determinada ação levou a um resultado ruim.

Dimensão Moral: Muitas emoções complexas, tais como culpa, remorso, orgulho, vergonha, admiração, constrangimento, etc; envolvem aspectos normativos e morais. Essas emoções são muitas vezes baseadas na internalização de um agente sobre alguma norma social. A norma social interiorizada por um agente torna-se um ideal do agente que o agente está empenhado em prosseguir e defender punindo possíveis infratores. Por exemplo, para se sentir culpado por algo não é suficiente que um agente acredite ser responsável pela violação de uma certa prescrição. Mas sim, achar que fez algo que não está de acordo com as normas que ele tem internalizadas.

Dimensão de responsabilidade: Algumas emoções complexas, como arrependimento, culpa, vergonha, remorso, repreensão, etc; baseiam-se em uma auto-atribuição de responsabilidade (por exemplo vergonha, culpa, remorso) e atribuição de responsabilidade à outros agentes (por exemplo, censura, admiração). Um agente sente admiração por alguém só se ele pensa que o outro foi responsável por algo que aconteceu, que está de acordo com seus ideais. Em particular, *responsabilidade completa* deve ser diferenciada da *responsabilidade parcial* ou *co-responsabilidade*. No primeiro caso, um agente i é dito ser totalmente responsável pela ocorrência de algum estado P se e somente se o agente provocou P independentemente do que o outros agentes fizeram. No segundo caso, o agente i é dito ser co-responsável pela ocorrência de P se, e somente se i provocou P juntamente com alguns outros agentes, isto é, P foi o efeito de uma ação conjunta de um grupo de agentes que inclui i . Esta distinção entre responsabilidade total e co-responsabilidade permite uma melhor compreensão da forma como a intensidade das emoções complexas devem ser calculadas. Por exemplo, o sentimento de culpa é mais intenso quando o agente acredita ser plenamente responsável por ter violado uma norma que internalizou ao invés de quando ele acredita ser co-responsável pela violação de uma norma.

Uma lógica para emoções complexas deve ser suficientemente expressiva, não só para caracterizar diferentes tipos de atitudes mentais dos agentes (crenças, desejos, objetivos, intenções), mas também para caracterizar o conceitos de responsabilidade, contrafactualidade, normas e ideais.

Nesse sentido, o modelo utiliza uma lógica STIT, que é uma lógica que suporta o raciocínio sobre as responsabilidades e poderes dos agentes, e

raciocínio contrafactual sobre ações e escolhas dos agentes. A semântica da STIT é baseada em uma estrutura de ramificação do tempo com momentos ordenados. Na lógica STIT o operador da forma $Stit_i\phi$ significa que um determinado agente i faz com que o estado ϕ seja verdadeiro não importando o que os outros agentes façam e o operador $Stit_C\phi$ representa o fato do grupo de agentes C fazerem com que o estado ϕ seja verdadeiro, não importando o que os agentes não pertencentes a C façam.

Assim sendo, adicionando-se operadores modais de crenças e objetivos com as construções padrão da lógica STIT, pode-se formalizar o conceito de arrependimento (*regret*) da seguinte forma:

$$Regret_i\phi \Leftrightarrow Bel_i\phi \wedge Bel_iX^- \neg Stit_{AGT \setminus \{i\}}X\phi \wedge Goal_i\neg\phi \quad (4.3)$$

Onde AGT denota o grupo incluindo todos os agentes, X expressa os fatos que serão verdadeiros no próximo momento e X^- expressa os fatos que eram verdadeiros no momento anterior (o símbolo $^-$ é um operador temporal que refere-se ao passado).

De acordo com a definição acima, o agente i arrepende-se por ϕ (denotado por $Regret_i\phi$) se e somente se i quer que ϕ seja falso (denotado por $Goal_i\neg\phi$), acredita que ϕ é verdadeiro (denotado por $Bel_i\phi$), e acredita que ele poderia ter agido de outra maneira para assegurar que ϕ fosse falso agora (denotado por $Bel_iX^- \neg Stit_{AGT \setminus \{i\}}X\phi$).

De uma forma mais prática, considerando uma situação em que a somente dois agentes i e j , isto é, $AGT = \{i, j\}$. O agente i decide estacionar seu carro em uma área proibida. O agente j (o policial) multa i em R\$ 100. O agente i arrepende-se por ter levado uma multa de R\$ 100 (denotado por $Regret_i\text{multa}$). Isto significa que, i quer que a *multa* seja falsa (denotada por $Goal_i\neg\text{multa}$), acredita que a *multa* é verdadeira (denotada por $Bel_i\text{multa}$), e acredita que poderia ter feito algo diferente (estacionar em outro local) e assegurar que a *multa* fosse falsa agora (denota por $Bel_iX^- \neg Stit_{AGT \setminus \{i\}}X\text{multa}$).

4.2.2 EBDI: Uma Arquitetura para Agentes Emocionais

Assim como o modelo proposto por (LORINI, 2008), esse modelo utiliza a arquitetura BDI como base para a inserção de emoções em agentes. Com a introdução de emoção primária e secundária na arquitetura BDI, o modelo, apresenta uma arquitetura genérica para um agente emocional, EBDI (acrônimo de *Emotions, Belief, Desire and Intention* - Emoções, Crença,

Desejo e Intenção), que pode mesclar várias teorias de emoções com o processo de raciocínio de um agente. Essa arquitetura implementa técnicas de raciocínio prático separadamente do mecanismo de emoção específico. A separação permite ligar modelos emocionais conforme a necessidade ou atualizar o motor de raciocínio do agente de forma independente (JIANG; VIDAL; HUHNS, 2007).

Nesse modelo, a arquitetura EBDI acrescenta a influência de emoções primárias e secundárias no processo de tomada de decisão de uma arquitetura BDI tradicional. As emoções primárias são utilizadas como o primeiro filtro para ajustar a prioridade de crenças, assim, permitindo aos agentes acelerarem a tomada de decisão. Por sua vez, as emoções secundárias são utilizados para refinar a decisão, caso o tempo permita.

Segundo o autor, para incorporar emoções em agentes, precisa-se resolver três problemas:

1. Como medir ou apresentar emoções?
2. Como as emoções afetam o processo de tomada de decisão?
3. Como atualizar o status de emoções?

Na maioria das vezes, as soluções dessas questões dependem da aplicação. A arquitetura EBDI combina esses 3 aspectos em uma arquitetura BDI com base no processo de raciocínio prático humano, deixando os detalhes em aberto para os desenvolvedores.

Na arquitetura EBDI, o processo de deliberação e raciocínio meio-fim é dividido em quatro componentes (Emoção, Desejo, Crença e Intenção) que são conectados através de algumas funções principais. Esses 4 componentes são explicados a seguir (JIANG; VIDAL; HUHNS, 2007).

- **Emoção:** São representadas como estados de consciência. Para esta componente, o método de representação de emoções, pode ser através de lógica de primeira ordem, lógica multidimensional, algum método de medição numérica ou outra abordagem.
- **Crença:** É geralmente definida como uma convicção da verdade de uma proposição, podendo ser adquiridas através da percepção, contemplação ou comunicação. Praticamente, as crenças são subjetivas para os seres humanos. O modelo BDI original obtém suas crenças a partir da função *see*, que sente o ambiente objetivamente. Na arquitetura EBDI, as crenças são influenciadas pelo estado emocional do agente e, em vez de adquirir crenças através apenas da percepção, foram adicionados métodos alternativos para obter crenças através da contemplação e comunicação.

- **Desejo:** Refere-se as opções que estão disponíveis para o agente ou os possíveis cursos de ações disponíveis para o agente. Desejos são obtidos através de uma *função de geração de opções*, com base nas crenças e intenções atuais.
- **Intenção:** As intenções desempenham um papel crucial no processo de raciocínio prático, porque elas tendem a levar à ação. São definidos 5 papéis importantes em relação ao raciocínio prático.
 - *Intenções que guiam o raciocínio meio-fim* - Como nos seres humanos, uma vez que um agente formula uma intenção, ele tentará alcançá-la e decidirá como irá conseguir isso. Se um determinado curso da ação falha ao tentar alcançar uma intenção, o agente normalmente tenta outros.
 - *Intenções persistentes* - Um agente não vai desistir de suas intenções até que ele acredite tê-las alcançado com sucesso, ou até que seja impossível alcançá-las, ou se a razão para a intenção não estiver mais presente.
 - *Emoções atuais influenciam a determinação das intenções* - Os agentes BDI determinam suas intenções com base em seus desejos e crenças. Algumas pesquisas mostram que a tomada de decisão pode, na verdade, ser conduzida por emoções. Essas pesquisas baseiam-se no fato de que pessoas com deficiências emocionais pequenas tem dificuldade em tomar decisões.
 - *Intenções restringem a deliberação futura* - um agente não vai receber opções que são incompatíveis com as suas intenções atuais.
 - *Intenções influenciam as emoções sobre as quais, juntamente com as crenças, o raciocínio prático futuro é baseado* - Uma vez que um agente adota uma intenção que possa planejar o futuro na suposição de que irá alcançar o objetivo. Assim, se houver uma crença que o agente não possa se beneficiar da intenção, o agente vai sentir-se triste, e essa emoção também irá influenciar futuras crenças e raciocínios.

A arquitetura EBDI é formalmente definida como: seja E o conjunto de todas as possíveis emoções, B é o conjunto de todas as crenças, D o conjunto de todos os desejos possíveis, e I o conjunto de todas as intenções possíveis. O estado de um agente EBDI em um dado momento é dada pelo seu conjunto atual de emoções, crenças, desejos e intenções. Estes componentes são ligados através das seguintes funções:

Funções de revisão de crenças: As crenças podem ser adquiridas a partir de percepção, contemplação ou comunicação, ao contrário do modelo BDI original, que usa apenas a percepção. São definidas 3 funções de revisão de crença que mapeiam as entradas dessas 3 áreas em crenças:

- *Função de revisão de crença através da percepção ($brf - see$):* gera as crenças candidatas a partir do ambiente:

$$brf - see : Env \rightarrow B_p \quad (4.4)$$

onde Env denota o ambiente, $B_p \subseteq B$ é o conjunto das possíveis crenças candidatas proveniente da percepção.

- *Função de revisão de crença através da Comunicação ($brf - msg$):* gera as crença candidatas a partir do conteúdo das mensagens de comunicação:

$$brf - msg : Cont \rightarrow B_m \quad (4.5)$$

onde $Cont$ denota o conteúdo das possíveis mensagens de comunicação, B_m é o conjunto das possíveis crenças candidatas proveniente da mensagem, e $B_m \subseteq B$.

- *Função de revisão de crença através da Contemplação ($brf - in$):* leva em consideração o estado de emoção e intenção atual, e revisa a base de crenças atual baseada nas crenças passadas e no conjunto de crenças candidatas provenientes do ambiente e das mensagens de comunicação. gera as crença candidatas a partir do conteúdo das mensagens de comunicação:

$$brf - in : E \times I \times (B \cup B_p \cup B_m) \rightarrow B \quad (4.6)$$

Funções de atualização de emoções: Uma vez que, nesse modelo são levadas em consideração as emoções primárias e emoções secundárias, são descritas 2 funções de atualização, uma para cada tipo de emoção:

- *Função de atualização de emoções primárias ($euf1$):*

$$euf1 : E \times I \times (B_p \cup B_m) \rightarrow E \quad (4.7)$$

As emoções primárias são aquelas que sentimos em primeiro lugar, como uma primeira resposta a uma situação. Assim, se estamos ameaçados, podemos sentir medo. Quando ouvimos falar de uma morte, podemos sentir tristeza. Elas aparecem antes do pensamento consciente e são, portanto, funções instintivas ou reativas do cérebro

humano. Quando o tempo é limitado e nós não temos tempo suficiente para pensar em algo claramente, as emoções primárias tornam-se extremamente úteis na tomada de decisões. Em agentes, podemos usar as emoções primárias para acelerar a tomada de decisão de forma semelhante.

- *Função de atualização de emoções secundárias (euf2):*

$$euf2 : E \times I \times B \rightarrow E \quad (4.8)$$

As emoções secundárias aparecem após as emoções primárias, podem ser causadas diretamente pelas emoções primárias, ou provenientes de cadeias mais complexas de pensamento. Por exemplo, quando o medo de uma ameaça se transforma em ira que alimenta o corpo para uma reação de luta. Para os agentes, as emoções secundárias vêm do resultado de novas deliberações e podem substituir as emoções primárias. Elas são usadas para refinar a tomada de decisão se o tempo permitir.

Função de geração de opção (options):

$$options : B \times I \rightarrow D \quad (4.9)$$

Função filtro (filter):

$$filter : E \times B \times D \times I \rightarrow I \quad (4.10)$$

Função plano (plan): As funções a seguir completam o processo de deliberação e geram as melhores opções-intenções, que podem ser ações ou estados mentais. Essas intenções guiam o raciocínio meio-fim que é representado pela função (plan):

$$plan : I \times A_c \rightarrow \pi \quad (4.11)$$

onde A_c é o conjunto de ações possíveis que um agente pode fazer, e π denota um plano que é uma sequência de ações

$$\pi = (\alpha_1, \dots, \alpha_n) \quad (4.12)$$

onde α_i é uma ação, e $\alpha_i \in A_c$

Função execução de plano (execute): essa função é utilizada para executar a sequência de ações produzidas pela função plano (plan). É representada por:

$$execute : \pi \rightarrow Env \quad (4.13)$$

Durante a execução, se π é vazio, ou a intenção atual é obter sucesso, ou o agente achar que a intenção atual é impossível de ser alcançada, então essa função será encerrada.

Uma característica importante desse modelo é que sua arquitetura implementa separadamente o processo de raciocínio prático e o mecanismo de emoção, permitindo uma fácil incorporação de novos modelos de emoções.

4.2.3 Formalização de emoções OCC com lógica BDI

Nesse modelo (ADAM et al., 2006) a arquitetura BDI é estendida para representar o maior número possível de emoções, presentes no modelo OCC, tentando também respeitar as definições originais dessas emoções da melhor forma possível. Esse framework é baseado no modelo lógico de crença, escolha, tempo e ação de (HERZIG; LONGIN, 2004). As noções de *escolha* e *intenção* foram suprimidas do modelo, no entanto, a lógica foi estendida com a adição de um operador modal de probabilidade, bem como os operadores de obrigação e desejo.

Nesse modelo os sentimentos de alegria (*joy*) e angústia (*distress*) podem ser representados como:

$$Joy_i\varphi := Bel_i\varphi \wedge Des_i\varphi \quad (4.14)$$

$$Distress_i\varphi := Bel_i\varphi \wedge Undes_i\varphi \quad (4.15)$$

onde $Bel_i\varphi$ representa a crença de i em φ , $Des_i\varphi$ representa que φ é desejável para i , e $Undes_i\varphi$ representa que φ é indesejável para i .

Esse modelo formaliza fielmente 20 tipos de emoções presentes no modelo OCC 4.1, porém ele não leva em consideração a intensidade das emoções nem a temporalidade das mesmas.

De uma maneira geral, os modelos descritos até agora apresentam ideias semelhantes para modelar as emoções em agentes, utilizando diversos tipos de lógicas e extensões para a criação de um modelo formal. É importante deixar claro que, nessa dissertação, a idéia principal não é criar uma nova lógica de emoções, mas sim propor um modelo de confiança baseado em emoções, ou seja, mostrar uma possível forma de unir dois modelos distintos de confiança e emoção. Assim, a lógica utilizada na parte emocional do modelo proposto não é tão complexa quanto as lógicas utilizadas pelos modelos apresentados nessa seção.

Para atingir nosso objetivo, nós consideramos a introdução de valores numéricos que representam a intensidade das emoções no modelo de confiança Regret 3.4. E, apesar de sua natureza qualitativa, as emoções são mapeadas em valores quantitativos que irão compor o cálculo de confiança e reputação. Assim, o modelo descrito a seguir será utilizado como base para a parte emocional do modelo TrustE, uma vez que, ele utiliza uma abordagem quantitativa para modelar as emoções.

4.2.4 Modelo de Steunebrink

Para formalizar os 22 tipos de emoções presentes no modelo OCC, Steunebrink, Dastani e Meyer (2009) identificaram ambiguidades existentes na estrutura lógica do OCC que dificultam a formalização das emoções. A partir dessa identificação, foram feitas alterações que resultaram em uma estrutura baseada em herança (Figura 4), suportada por uma nova estrutura lógica, e novas especificações dos tipos de emoções. Partindo dessas alterações foi proposto uma lógica para formalização das emoções em agentes (STEUNEBRINK; DASTANI; MEYER, 2007a); (STEUNEBRINK; DASTANI; MEYER, 2012), e iniciado um trabalho para quantificar as intensidades das emoções (STEUNEBRINK; DASTANI; MEYER, 2007b), através da criação de funções que forneçam, como resultado, a intensidade das emoções.

4.2.4.1 Nova hierarquia de emoções

Uma das principais alterações feitas na estrutura do modelo OCC foi o fato da hierarquia ter uma herança explícita, com rótulos em cada ponto da hierarquia e as condições de cada nodo filho serem um superconjunto dos seus nodos pais. A partir dessa nova estrutura foram criadas novas especificações dos tipos de emoções, conforme a figura 5.

Apesar das modificações realizadas no modelo OCC, a hierarquia do modelo continua sendo dividida em 3 ramos e as emoções *satisfeito*, *insatisfeito*, *aprovação*, *desaprovação*, *gosto* e *desgosto* continuam sendo a base para todas as outras emoções assim como no modelo original OCC.

Figura 5 – Especificações dos tipos de emoções do modelo OCC modificado (STEUNEBRINK; DASTANI; MEYER, 2009) [tradução nossa]

positivo e *negativo* são reações atrativas ou aversivas (em relação a "alguma coisa").
satisfeito é ser *positivo* em relação a consequência (de um evento).
insatisfeito é ser *negativo* em relação a consequência (de um evento).
esperança é ser *positivo* em relação a uma possível consequência (de um evento).
medo é ser *negativo* em relação a uma possível consequência (de um evento).
alegria é ser *positivo* em relação a consequência atual (de um evento).
angústia é ser *negativo* em relação a consequência atual (de um evento).
satisfação é a *alegria* em relação a confirmação de uma possível consequência desejável.
medo-confirmado é a *angústia* em relação a confirmação de uma possível consequência indesejável.
alívio é a *alegria* em relação a não confirmação de uma possível consequência indesejável.
desapontamento é a *angústia* em relação a não confirmação de uma possível consequência desejável.
feliz-por é a *alegria* em relação a consequência (de um evento) que presume-se ser desejável para alguém.
ressentimento é a *angústia* em relação a consequência (de um evento) que presume-se ser desejável para alguém.
regozijo é a *alegria* em relação a consequência (de um evento) que presume-se ser indesejável para alguém.
pena é a *angústia* em relação a consequência (de um evento) que presume-se ser indesejável para alguém.
aprovação é ser *positivo* em relação a ação (de um agente).
desaprovação é ser *negativo* em relação a ação (de um agente).
orgulho é a *aprovação* da sua própria ação.
vergonha é a *desaprovação* da sua própria ação.
admiração é a *aprovação* da ação de outro agente.
reprovação é a *desaprovação* da ação de outro agente.
gratificação é o *orgulho* em relação a uma ação e a *alegria* em relação a consequência relacionada.
remorso é o *vergonha* em relação a uma ação e a *angústia* em relação a consequência relacionada.
gratidão é a *admiração* em relação a uma ação e a *alegria* em relação a consequência relacionada.
raiva é a *reprovação* em relação a uma ação e a *angústia* em relação a consequência relacionada.
gosto é ser *positivo* em relação a um aspecto (de um objeto).
desgosto é ser *negativo* em relação a um aspecto (de um objeto).
amor é o *gosto* por um aspecto conhecido (de um objeto).
ódio é o *desgosto* por um aspecto conhecido (de um objeto).
interesse é o *gosto* por um aspecto desconhecido (de um objeto).
desinteresse é o *desgosto* por um aspecto desconhecido (de um objeto).

emoções são desencadeadas. A seguir são apresentadas as formalizações de 4 tipos de emoções segundo o modelo (STEUNEBRINK; DASTANI; MEYER, 2007a), onde a é um agente, π um plano e k um objetivo.

- $ALEGRIA_a(k)$ - representa a alegria do agente a ao cumprir um certo objetivo k ;
- $ANGUSTIA_a(k)$ - representa a angustia do agente a em não cumprir um objetivo k ;
- $ESPERANCA_a(\pi, k)$ - representa a esperança que o agente a tem de que ao executar o plano π alcançará o objetivo k ;
- $MEDO_a(\pi, -k)$ - representa o medo que o agente a tem de que ao executar o plano π não alcançará o objetivo k .

Uma vez que esse trabalho especifica as emoções em nível de linguagem de agentes, as formalizações das emoções levam em consideração os planos e objetivos dos agentes. No modelo proposto esses aspectos não serão abordados uma vez que o modelo está em um nível mais abstrato, em que a linguagem dos agentes não é relevante. Assim sendo, a descrição das emoções precisa sofrer algumas modificações para adequar-se ao modelo proposto, conforme é demonstrado a seguir, onde a , b são agentes, X um evento e Y uma ação.

- $SATISFEITO_a(X)$ - o agente a tem uma *reação positiva* em relação a consequência de um evento X ;
- $INSATISFEITO_a(X)$ - o agente a tem uma *reação negativa* em relação a consequência de um evento X ;
- $APROVACAO_a(Y)$ - o agente a tem uma *reação positiva* em relação a ação Y de algum agente.
- $DESAPROVACAO_a(Y)$ - o agente a tem uma *reação negativa* em relação a ação Y de algum agente.
- $ALEGRIA_a(X)$ - o agente a está *satisfeito* a respeito da atual consequência de um evento X ;
- $ANGUSTIA_a(X)$ - o agente a está *insatisfeito* a respeito da atual consequência de um evento X ;
- $ESPERANCA_a(X)$ - o agente a está *satisfeito* sobre uma possível consequência de um evento X ;
- $MEDO_a(X)$ - o agente a está *insatisfeito* sobre uma possível consequência de um evento X .
- $ORGULHO_a(Y, b)$ - o agente a *aprovou* a sua própria ação Y ;
- $VERGONHA_a(Y, b)$ - o agente a *desaprovou* a sua própria ação Y ;
- $ADMIRACAO_a(Y, b)$ - o agente a *aprovou* a ação Y do agente b ;
- $REPROVACAO_a(Y, b)$ - o agente a *desaprovou* a ação Y do agente b ;

A seguir será introduzido o conceito de intensidade das emoções em agentes, que é um fator importantíssimo para o modelo proposto dessa dissertação, pois é através do valor das intensidades que será feita a inserção das emoções no modelo de confiança.

4.2.4.3 Intensidade das Emoções

Conforme dito anteriormente, no modelo OCC, os aspectos quantitativos das emoções são descritos em termos de *potencialidades*, *limiares* e *intensidades*. O termo *limiar* refere-se ao valor mínimo necessário para que a emoção possua alguma intensidade, e consequentemente, interfira no comportamento do agente. Por sua vez a *potencialidade* de uma emoção depende de uma lista de variáveis que afetam a sua intensidade. Quanto maior a *potencialidade* de uma emoção, maior as chances dela ocorrer.

A intensidade de uma emoção é definida subtraindo-se o limiar de sua potencialidade, podendo ser 0 (zero) caso o limiar seja maior que seu potencial. Entretanto o modelo OCC não especifica como calcula os limiares das emoções, porém acredita-se que eles dependam de variáveis globais que indicam o humor do agente (STEUNEBRINK; DASTANI; MEYER, 2007b). Por exemplo, se um agente está de bom humor, os limiares das emoções negativas aumentam, causando uma diminuição na intensidade dessa emoções.

Quando uma condição necessária para disparar uma emoção ocorre, mas o seu potencial está abaixo do seu limiar, um agente pode reconhecer que essa emoção foi desencadeada porém ela não o afetará. Por exemplo, "o humor de um agente estava tão bom que mesmo ele tendo praticado uma ação ruim, ele não foi afetado pela vergonha". Diferentemente da abordagem adotada nos trabalhos descritos anteriormente em que a emoção, por ser tratada somente de forma qualitativa, se restringe em afetar ou não o comportamento do agente, ao tratar as emoções de forma quantitativa, através da sua intensidade, é possível alterar o grau (intensidade) em que uma determinada emoção afetará o comportamento do agente. Esse fato torna o modelo mais dinâmico pois possibilita uma maior quantidade de alterações no comportamento dos agentes através dos diferentes valores de intensidade que cada emoção pode assumir.

A *intensidade* de uma emoção é calculada a partir da sua *função de intensidade*, que é composta pela *função de potencialidade* e *função de limiar*. A maneira como essas funções são calculadas é dependente da aplicação, porém, de um modo geral, a *função de intensidade* de uma emoção, denotada por $I(P(E), L(E), t)$, que por simplificação será representada por $I(E)$, pode ser expressa como:

$$I(P(E), L(E), t) \rightarrow \mathbb{R}^+ \quad (4.16)$$

onde $P(E)$ é a *função de potencialidade* da emoção E , $L(E)$ a *função de limiar* e t o tempo corrente. Como resultado, a função retorna um real

positivo, incluindo 0 (zero), representando quantitativamente a intensidade da emoção. O valor de $I(E)$ persiste ao longo do tempo, e tende a diminuir com o passar dele, diferentemente dos valores de $P(E)$ e $L(E)$ que são recalculados a cada vez que E é disparada novamente, portanto esses valores não persistem com o passar do tempo. Para suportar a temporalidade do valor de $I(E)$ cada agente deve possuir uma memória (MEA) que armazene os valores de cada uma das emoções sentidas por ele (STEUNEBRINK; DASTANI; MEYER, 2007b).

No capítulo seguinte será apresentado e explicado o modelo proposto nesse trabalho.

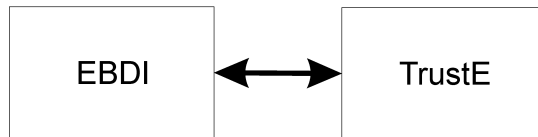
5 MODELO PROPOSTO

O modelo TrustE, proposta deste trabalho, agrega emoções ao cálculo da confiança e da reputação em agentes. Todos os modelos encontrados na literatura fazem uso de análise algébrica para determinar o ato de confiar, o que torna tais modelos essencialmente matemáticos e desconectados do histórico de agente. A proposta do modelo TrustE é permear o modelo de confiança com avaliações oriundas de raciocínio simbólico, tornando o ato de confiar mais dinâmico e dependente do histórico do agente. Isso é possível através da incorporação de emoções em modelos de confiança.

A incorporação de emoções pode ser feita de várias formas, contudo, optou-se por incorporar fatores emocionais ao modelo Regret 3.4, baseados nas *intensidades das emoções* do agente. Assim, os pesos ξ e ω (presentes nas equações 3.31, 3.33, 3.35 e 3.37) definidos pelo programador no modelo Regret são substituídos por pesos que são calculados em tempo de execução e decorrem das experiências vividas pelo agente.

As emoções são incluídas no modelo através de suas *intensidades*, ou seja, por um valor numérico que representa a força da emoção em um determinado momento t . Isso permite ao modelo aceitar cada uma das 22 emoções existentes no modelo OCC (apresentado na seção 4.1), exigindo apenas que os aspectos qualitativos e quantitativos de cada uma das emoções sejam modelados. Os aspectos qualitativos podem seguir as mesmas ideias presentes nos modelos citados na Seção 4.2 ou serem adaptados de acordo com a necessidade de cada sistema. Os aspectos quantitativos são obtidos a partir de cada emoção, sendo calculados através de uma *função de intensidade*, e podem seguir os conceitos apresentados em (STEUNEBRINK; DASTANI; MEYER, 2007b) ou serem reformulados de acordo com a necessidade do desenvolvedor. Essa flexibilidade é uma característica fundamental do modelo.

Figura 6 – Interface entre arquitetura EBDI e o modelo TrustE



O modelo TrustE é responsável por disparar as emoções que dependem da análise dos resultados do cálculo de confiança pelos agentes.

No entanto, conforme dito na Seção 1.1, os agentes podem sentir emoções externas ao modelo. Essas emoções podem ser expressas por uma arquitetura EBDI, por exemplo, que seria responsável por armazená-las e gerenciá-las. Assim, o TrustE utilizaria suas emoções internas em conjunto com as emoções da arquitetura EBDI para calcular a confiança. Para que isso seja possível, basta definir uma interface entre o TrustE e a arquitetura EBDI (figura 6) para que haja uma troca de informações entre ambos, permitindo que as emoções modeladas na arquitetura EBDI também façam parte do cálculo da confiança.

5.1 FORMALIZAÇÃO DAS EMOÇÕES NO TRUSTE

Na implementação do TrustE utilizou-se oito emoções presentes no modelo OCC 4.1, divididas em dois grupos:

- Emoções disparadas por eventos: *esperança, medo, alegria e angústia*.
- Emoções disparadas por ações: *orgulho, vergonha, admiração e reprovação*.

O primeiro grupo representa as emoções que são induzidas a partir da avaliação de um evento com respeito aos objetivos do agente, ou seja, se a ocorrência do evento ajudou o agente de uma maneira geral. O segundo grupo envolve as emoções que são baseadas no julgamento de uma ação realizada por um determinado agente.

Essas emoções também são classificadas em emoções positivas (*EP*) e emoções negativas (*EN*):

- Emoções positivas (*EP*): *alegria, esperança, orgulho, admiração*
- Emoções negativa (*EN*): *angústia, medo, vergonha, reprovação*

A modelagem formal dessas emoções foram simplificadas para adequar-se ao modelo e tornar o cálculo das *intensidades* de cada uma dessas emoções mais simples. A seguir será apresentado a formalização dessas emoções levando em conta alguns ajustes realizados no modelo OCC, apresentados na seção 4.2.4.

Sejam a e b agentes, X um evento, Y uma ação, *positivo* uma reação positiva em relação a alguma coisa, e *negativo* uma reação negativa em relação a alguma coisa.

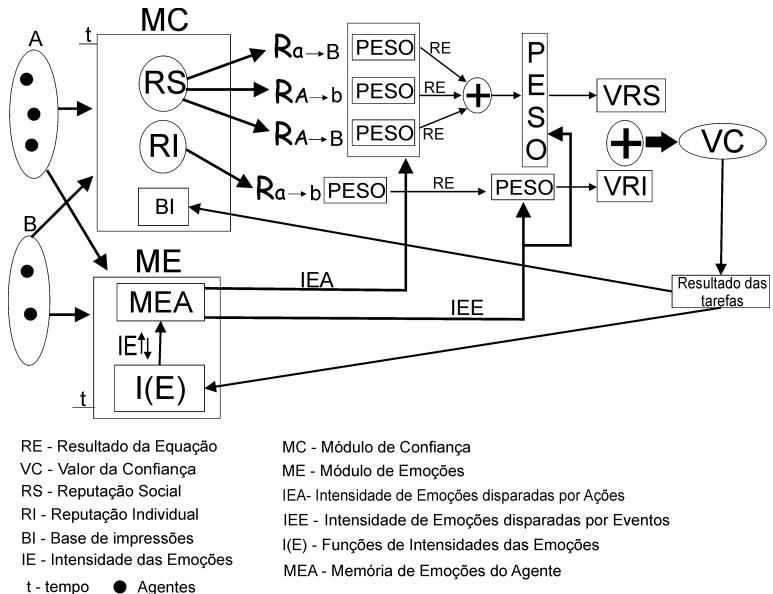
- $Acao_a(Y)$: uma ação Y realizada por a .

- $Provavel(X)$: um evento X que possivelmente irá ocorrer.
- $Atual(X)$: um evento X que está ocorrendo.
- $Satisfeito_a(X)$: a é positivo sobre a consequência de X .
- $Insatisfeito_a(X)$: a é negativo sobre a consequência de X .
- $Aprovacao_a(Y)$: a é positivo sobre a ação Y de algum agente.
- $Desaprovacao_a(Y)$: a é negativo sobre a ação Y de algum agente.

5.1.1 Emoções disparadas por eventos

Essas emoções são sentidas pelos agentes conforme a ocorrência de um determinado evento e não dependem das ações realizadas pelos outros agentes. Por essa razão elas são introduzidas no modelo TrustE (ver figura 7) através dos valores das funções IEE . Os itens mostrados na figura 7 serão descritos em detalhes ao longo do texto.

Figura 7 – Modelo TrustE



As emoções disparadas por eventos são definidas como:

$$Esperanca_a(X) \leftrightarrow Satisfeito_a(X) \wedge Provavel(X) \quad (5.1)$$

onde os símbolos ‘ \leftrightarrow ’ e ‘ \wedge ’ tem o significado usual. Um agente sente esperança se ele está satisfeito em relação a provável ocorrência de um evento. *Exemplo* : ‘O agente tem esperança que o produto comprado seja de boa qualidade’.

$$Medo_a(X) \leftrightarrow Insatisfeito_a(X) \wedge Provavel(X) \quad (5.2)$$

Um agente sente medo se ele está insatisfeito em relação a provável ocorrência de um evento. *Exemplo* : ‘O agente estava com medo de não receber o produto comprado’.

$$Alegria_a(X) \leftrightarrow Satisfeito_a(X) \wedge Atual(X) \quad (5.3)$$

Um agente sente alegria quando está satisfeito em relação a um evento atual. *Exemplo* : ‘O agente estava feliz por ter recebido o produto antes do prazo’.

$$Angustia_a(X) \leftrightarrow Insatisfeito_a(X) \wedge Atual(X) \quad (5.4)$$

Um agente sente angústia quando está insatisfeito em relação a um evento atual. *Exemplo* : ‘O agente está angustiado por ter recebido um produto de má qualidade’.

5.1.2 Emoções disparadas por ações

Essas emoções são sentidas pelos agentes como resultado de ações realizadas pelos outros agentes. Por essa razão elas são introduzidas no modelo TrustE (ver figura 7) através dos valores das funções *IEA*.

As emoções disparadas por ações são definidas como:

$$Orgulho_a(Y,a) \leftrightarrow Aprovacao_a(Y) \wedge Acao_a(Y) \quad (5.5)$$

Um agente sente orgulho se ele aprova sua própria ação. *Exemplo* : ‘O agente estava orgulhoso de ter vendido um produto de alta qualidade’.

$$Vergonha_a(Y,a) \leftrightarrow Desaprovacao_a(Y) \wedge Acao_a(Y) \quad (5.6)$$

Um agente sente vergonha se ele desaprova sua própria ação. *Exemplo* : ‘O agente sentiu-se envergonhado de ter entregue o produto depois do prazo’.

$$Admiracao_a(Y,b) \leftrightarrow Aprovacao_a(Y) \wedge Acao_b(Y) \quad (5.7)$$

Um agente sente admiração por outro agente, se ele aprova a ação desse agente. *Exemplo* : ‘O agente estava admirado com a honestidade do vendedor’.

$$Reprovacao_a(Y,b) \leftrightarrow Desaprovacao_a(Y) \wedge Acao_b(Y) \quad (5.8)$$

Um agente sente reprovação por outro agente, se ele desaprova a ação deste agente. *Exemplo* : ‘O agente reprovou o fato do vendedor não responder suas dúvidas’.

Cada uma dessas oito fórmulas lógicas representa o gatilho (“trigger”) de cada uma das emoções presentes no modelo. Esses gatilhos definem quando uma emoção é disparada (sentida) por um determinado agente.

A Seção seguinte apresenta as *funções de intensidade* utilizadas no modelo TrustE.

5.2 INTENSIDADE DAS EMOÇÕES NO TRUSTE

Conforme dito anteriormente, a proposta do modelo TrustE é agregar emoções ao cálculo da confiança e da reputação em agentes. No TrustE isso é feito através da incorporação das intensidades das emoções no modelo de confiança Regret (apresentado na seção 3.4). No modelo TrustE as intensidades das emoções persistem ao longo do tempo e tendem a estabilizar-se com o passar dele, ou seja, a intensidade de cada uma das emoções aumenta ou diminui gradativamente, ao longo do tempo, até assumir seu valor padrão. Para suportar essa temporalidade cada agente possui uma memória (MEA) que armazena os valores das emoções sentidas por ele (ver figura 7). A MEA será descrita na próxima seção 5.2.1.

Cada uma das emoções do TrustE poderia ser calculada através de uma função distinta, no entanto, por uma questão de simplificação e para manter o escopo desse trabalho, optou-se por formular uma única equação (5.9) capaz de modelar as intensidades das emoções.

$$I_a(E,t) = \max(0, P_a(E,t) - L_a(E,t)) \quad (5.9)$$

$I_a(E,t)$ é a *função de intensidade* da emoção E do agente a no tempo t ; E é uma das oito emoções do modelo TrustE; $P_a(E,t)$ é a *potencialidade* de E ; $L_a(E,t)$ o *limiar* de E e t o tempo corrente. Conforme descrito no

modelo OCC, a intensidade de uma emoção é definida subtraindo-se o limiar de sua potencialidade. Uma emoção E terá uma intensidade alta quando $P_a(E, t) \geq L_a(E, t)$, baixa quando $P_a(E, t) = L_a(E, t)$ e 0 (zero) quando $P_a(E, t) \leq L_a(E, t)$. Essa função é recalculada cada vez que o agente a sente a emoção E , e seu valor é utilizado na função (5.14) para atualizar a MEA do agente.

A função de *potencialidade* $P_a(E, t)$ é calculada através da fórmula:

$$P_a(E, t) = \frac{\sum_{i=1}^N VI_i}{N} \quad (5.10)$$

onde $VI \in [0, 1]$ são os valores das variáveis que afetam a *intensidade* da emoção E (segundo o modelo OCC, seção 4.1), e N é o número de variáveis.

O *limiar* $L_a(E, t)$, que representa o ‘perfil emocional’ do agente, é calculado através da fórmula:

$$L_a(E, t) = \begin{cases} 1 - H(a, t) & \text{se } E \in EP \\ H(a, t) & \text{caso } E \in EN \end{cases} \quad (5.11)$$

onde $H(a, t) \in [0, 1]$ é a função que representa o humor do agente a no tempo t . Para agentes de ‘bom humor’ a função $H(a, t)$ retorna valores maiores, assim o valor de $L_a(E, t)$ será menor quando E for uma emoção positiva (EP) e maior quando E for negativa (EN). Já para agentes de ‘mau humor’ a função $H(a, t)$ retorna valores menores, tornando os *limiares* das EP maiores e das EN menores. Assim, um agente de ‘bom humor’ vai sentir emoções positivas com maior facilidade do que um agente de ‘mau humor’, da mesma forma que um agente de ‘mau humor’ vai sentir emoções negativas mais facilmente do que um agente de ‘bom humor’.

Para exemplificar essa função, vamos levar em consideração que a função $H(a, t)$ retorna valores aleatórios no intervalo $[0.5, 1]$ para agentes de bom humor e valores aleatórios no intervalo $[0, 0.5]$ para agentes de mau humor. Supondo que a seja um agente com bom humor, $t = 3$, $E = \text{alegria}$ e $H(a, 3) = 0.8$, e dado que $\text{alegria} \in EP$, o *limiar* de *alegria* é calculado através da função $L_a(\text{alegria}, 3) = 1 - H(a, 3) = 1 - 0.8 = 0.2$. Como 0.2 é um valor baixo de limiar, quando o agente sentir alegria, sua intensidade tende a ser alta. Agora, imaginando outro cenário, temos que b é um agente de mau humor, $E = \text{esperança}$ e $H(b, 3) = 0.1$, assim teríamos que $L_b(\text{esperança}, 3) = 1 - H(b, 3) = 1 - 0.1 = 0.9$ já que $\text{esperança} \in EP$. Logo, como 0.9 é um valor alto de *limiar*, se b sentir esperança, sua intensidade será muita baixa ou talvez nula.

Uma vez que os valores de $P_a(E, t)$ e $L_a(E, t)$ estão no intervalo $[0, 1]$, a função $I_a(E, t)$ também está no intervalo $[0, 1]$.

5.2.1 Memória de emoções dos agentes

Cada agente possui uma memória (MEA) que é responsável por suportar a temporalidade das intensidades das emoções vivenciadas por eles. Essa memória é atualizada a cada momento que o agente sente uma emoção e as intensidades dessas emoções tendem a estabilizar-se com o passar do tempo. Para suportar essas características foi definida uma função $IP_a(E, t)$ que representa a intensidade padrão da emoção E do agente a no tempo t ; uma função $M_a(E)$ que incrementa ou decrementa a intensidade da emoção E , na MEA do agente a , com o passar do tempo até que ela atinja o valor de $IP_a(E, t)$; e uma função $AI_a(E, t)$ que atualiza a intensidade da emoção na MEA cada vez que E é sentida pelo agente. Na definição dessas funções a MEA do agente a é denotada por mea_a e a variável que representa o valor atual da intensidade de E presente na mea_a é expressa por $IMEA_a^E$.

A função $IP_a(E)$ é definida como:

$$IP_a(E, t) = vp \quad (5.12)$$

onde $vp \in [0, 1]$ representa o valor padrão da intensidade da emoção E do agente a no tempo t . O valor de vp depende do perfil emocional do agente.

A função $M_a(E)$ é definida como:

$$M_a(E, t) = \begin{cases} \max(IP_a(E, t), (IMEA_a^E - p)) & \text{se } IMEA_a^E > IP_a(E, t) \\ \min(IP_a(E, t), (IMEA_a^E + p)) & \text{se } IMEA_a^E < IP_a(E, t) \\ IP_a(E, t) & \text{se } IMEA_a^E = IP_a(E, t) \end{cases} \quad (5.13)$$

onde $p \in (0, 0.2]$, representa o fator de atualização da emoção a cada tempo t . Essa função incrementa ou decrementa o valor de $IMEA_a^E$ até que a emoção E atinja sua intensidade padrão, expressa por $IP_a(E, t)$. O valor de p pode ser escolhido de acordo com a necessidade do ambiente ou do humor do agente. No caso de agentes com humor inconstante, pode-se dar um valor maior a p , visto que pessoas que mudam facilmente de humor tendem a mudar mais rapidamente suas emoções. Para agentes com humor constante o valor de p deve ser menor, pois esses levarão um tempo maior para mudar as intensidades de suas emoções.

Para dar uma demonstração de utilização dessa função vamos supor que $E = angustia$, $t = 5$, $IP_a(angustia, 5) = 0.5$, $p = 0.2$ e $IMEA_a^{angustia} = 0.45$. Dado esses valores, temos que $IMEA_a^{angustia} < IP_a(angustia, 5)$, e calculamos $M_a(angustia)$ através da equação $\min(0.5, (0.45 + 0.2)) = \min(0.5, 0.65) = 0.5$. Assim o valor da intensidade da *angustia* na mea_a é

atualizado para 0.5.

Como as intensidades das emoções são recalculadas a cada vez que um agente sente uma emoção, e os valores das emoções persistem ao longo do tempo, a função $AI_a(E, t)$ é a responsável por atualizar o valor da intensidade da emoção E na mea_a sempre que o agente sentir E .

$$AI_a(E, t) = IMEA_a^E + \mu \cdot I_a(E, t) \quad (5.14)$$

onde $\mu = (1 - IMEA_a^E)$. A função agrega ao valor $IMEA_a^E$ uma fração μ da intensidade de E vivenciada pelo agente a no tempo t . O resultado dessa função será o novo valor da intensidade da emoção E presente na mea_a , denotado por $IMEA_a^E$. Por exemplo, sendo $E = admiracao_b$, $IMEA_a^{admiracao_b} = 0.6$, $t = 12$ e $I_a(admiracao_b, 12) = 0.8$. Temos que $\mu = (1 - IMEA_a^{admiracao_b}) = 1 - 0.6 = 0.4$. Utilizando esse valor na equação $AI_a(admiracao_b, 12)$ temos que $IMEA_a^{admiracao_b} + \mu \cdot I_a(admiracao_b, 12) = 0.6 + (0.4) \cdot (0.8) = 0.92$. Assim a intensidade da emoção $admiracao_b$ na mea_a é atualizada para 0.92.

Todas essas funções são dependentes de aplicação, e podem ser modificadas conforme a necessidade do sistema. Além dessas funções, a MEA também é responsável por armazenar as fórmulas lógicas (gatilhos) apresentadas na Seção 5.1.

A seguir será apresentado e explicado o funcionamento do modelo TrustE.

5.3 O MODELO TRUSTE

O modelo TrustE é um modelo de confiança baseado em emoções, e seu propósito principal é tornar o processo de calcular a confiança entre dois agentes mais subjetivo, tentando afastar a utilização de valores que não estão relacionadas ao “emocional” dos agentes. Para isso, ao invés de utilizar valores que não possuem relações com o passado “emocional” do agente, como os utilizados no modelo REGRET (os pesos ξ e ω das equações (3.31), (3.33), (3.35) e (3.37)), serão utilizados valores baseados nas *intensidades das emoções* presentes na MEA dos agentes.

No modelo Regret os valores de ω são ajustados de acordo com a proximidade entre o agente avaliador e o agente avaliado, ou podem refletir a credibilidade ou relações hierárquicas dentro do grupo, ou seja, quanto maior a credibilidade de um agente maior será o peso dado a sua informação. Já os valores de ξ são usados para dar uma maior ou menor importância a cada uma das quatro fontes de opinião ($R_{a \rightarrow b}(\varphi)$, $R_{a \rightarrow B}(\varphi)$, $R_{A \rightarrow b}(\varphi)$ e $R_{A \rightarrow B}(\varphi)$)

presentes na equação (3.37)). Esses valores não são necessariamente estáticos e podem mudar ao longo do tempo de acordo com o estado mental do agente, seguindo uma regra definida pelo desenvolvedor.

A equação (3.27) que calcula a Reputação Individual (RI) no modelo Regret foi usada no modelo TrustE na sua forma original. Essa equação não sofreu alterações pois acredita-se que a inserção de emoções nessa equação não traria mudanças significativas ao modelo, além de que, a proposta de fornecer valores maiores para as avaliações dadas recentemente é uma boa abordagem para o agente avaliar, de forma consistente, suas interações diretas com outros agentes. As equações que calculam a Reputação Social (RS) tiveram seus pesos ω alterados por “símbolos emocionais” que representam a intensidade das emoções disparadas por ações (*IEA*). Escolheu-se esse tipo de emoção pois a RS está ligada diretamente a opiniões de outros agentes, e através das *IEA* é possível que um determinado agente consiga avaliar os outros levando em consideração as emoções relacionadas a cada um deles. Por sua vez a equação geral (3.37) teve os pesos ξ alterados por “símbolos emocionais” que representam a intensidade das emoções baseada em eventos (*IEE*). Foi utilizado esse tipo de emoção na equação geral do Regret por considerarmos que essa equação está ligada diretamente ao estado emocional do agente, pois é através dela que o agente pode dar um peso maior a RI ou a RS. Assim, quando um agente estiver *angustiado*, por exemplo, ele pode dar uma importância maior a RI.

Os três valores da RS ($R_{a \rightarrow B}(\varphi)$, $R_{A \rightarrow b}(\varphi)$ e $R_{A \rightarrow B}(\varphi)$) da equação geral são referentes às informações oriundas de outros agentes. No modelo TrustE esses três valores foram unidos, através de uma média simples, formando um único valor. Assim a equação geral do TrustE (5.19) possui dois valores, sendo um baseado nas opiniões do próprio agente e outro baseado nas opiniões de terceiros.

Retomando as equações do modelo Regret e substituindo os pesos ω e ξ pelos valores *IEA* e *IEE* temos as seguintes equações:

$$R_{a \rightarrow b}(\varphi) = \sum_{t_i \in IDB_p^a} \rho(t, t_i) \cdot W_i \quad (5.15)$$

onde $\rho(t, t_i) = \frac{f(t_i, t)}{\sum_{t_j \in IDB_p^a} f(t_j, t)}$, e $f(t_i, t)$ é uma equação em função do tempo que atribui maiores valores à valores mais próximos de t (um exemplo dessa função pode ser $f(t_i, t) = \frac{t_i}{t}$), t é o tempo corrente e $R_{a \rightarrow b}(\varphi) \in [-1, 1]$. Essa equação não sofreu nenhuma alteração no modelo TrustE.

$$R_{a \rightarrow B}(\varphi) = \sum_{b_i \in B} IEA^{ab_i} \cdot R_{a \rightarrow b_i}(\varphi) \quad (5.16)$$

onde $\sum_{b_i \in B} IEA^{ab_i} = 1$ e $R_{a \rightarrow B}(\varphi) \in [-1, 1]$.

$$R_{A \rightarrow b}(\varphi) = \sum_{a_i \in A} IEA^{a_i b} \cdot R_{a_i \rightarrow b}(\varphi) \quad (5.17)$$

onde $\sum_{a_i \in A} IEA^{a_i b} = 1$ e $R_{A \rightarrow b}(\varphi) \in [-1, 1]$

$$R_{A \rightarrow B}(\varphi) = \sum_{a_i \in A} IEA^{a_i B} \cdot R_{a_i \rightarrow B}(\varphi) \quad (5.18)$$

onde $\sum_{a_i \in A} IEA^{a_i B} = 1$ e $R_{A \rightarrow B}(\varphi) \in [-1, 1]$

$$SR_{a \rightarrow b}(\varphi) = IEE_a^{RI} \cdot R_{a \rightarrow b}(\varphi) + IEE_a^{RS} \cdot R_{RS} \quad (5.19)$$

onde $IEE_a^{RI} + IEE_a^{RS} = 1$ e $R_{RS} = \left(\frac{(R_{a \rightarrow B}(\varphi) + R_{A \rightarrow b}(\varphi) + R_{A \rightarrow B}(\varphi))}{3} \right)$.

A figura 7 exemplifica a estrutura do modelo TrustE que é um modelo híbrido de confiança (MC) e emoções (ME). O módulo de confiança é o responsável pelo cálculo da RI (equação (5.15)) e da RS (equações (5.16), (5.17), (5.18)), que juntas irão resultar no valor final da confiança, denotado por VFC, expressa pela equação (5.19). O módulo ME contém as funções das intensidades das emoções $I(E)$ e as memórias das emoções dos agentes (MEA) (ver Seção 5.2).

Os valores das IEE dependem das quatro emoções disparadas por eventos (alegria, angústia, medo e esperança) e os valores das IEA dependem das emoções disparadas por ações (orgulho, vergonha, admiração e reprovação). Como pode ser observado na figura 7 as IEE são utilizadas como os dois pesos da equação geral (5.19), enquanto os pesos IEA^{ab_i} , $IEA^{a_i b}$ e $IEA^{a_i B}$ são utilizados nas três equações da RS. Para o peso IEE_a^{RI} serão levados em consideração as emoções *esperança* e *angústia* do agente a enquanto para IEE_a^{RS} as emoções *alegria* e *medo*. Já para os pesos da RS serão consideradas as emoções *reprovação* para IEA^{ab_i} , *aprovação* para $IEA^{a_i b}$ e *orgulho* e *vergonha* para $IEA^{a_i B}$.

Vale lembrar que a distribuição dessas emoções entre os pesos IEE e IEA pode variar de acordo com a necessidade da aplicação, e essa configuração, aqui descrita, representa uma das várias possibilidades.

A seguir são apresentadas as equações que representam os valores de IEE_a^{RI} , IEE_a^{RS} , IEA^{ab_i} , $IEA^{a_i b}$ e $IEA^{a_i B}$.

Seja X_a^{RI} e X_a^{RS} os valores não normalizados de IEE_a^{RI} e IEE_a^{RS} respectivamente, então temos que:

$$X_a^{RI} = \frac{IMEA_a^{esperanca} + (1 - IMEA_a^{angustia})}{2} \quad (5.20)$$

$$X_a^{RS} = \frac{IMEA_a^{alegria} + (1 - IMEA_a^{medo})}{2} \quad (5.21)$$

Para permitir que o valor final da confiança (VFC) esteja no intervalo $[-1, 1]$, precisa-se que $IEE_a^{RI} + IEE_a^{RS} = 1$, logo X_a^{RI} e X_a^{RS} precisam ser normalizados. Assim, finalmente temos os valores de IEE_a^{RI} e IEE_a^{RS} dados pelas equações:

$$IEE_a^{RI} = \frac{X_a^{RI}}{(X_a^{RI} + X_a^{RS})} \quad (5.22)$$

$$IEE_a^{RS} = 1 - IEE_a^{RI} \quad (5.23)$$

onde $X_a^{RI} + X_a^{RS} > 0$

Da mesma forma que os valores das IEE , os valores das IEA precisam ser normalizados. Assim, seja Y^{abi} , Y^{aib} e Y^{a_iB} os valores não normalizados de IEA^{abi} , IEA^{aib} e IEA^{a_iB} respectivamnete, então temos que:

$$Y^{abi} = 1 - reprovacao_{a \rightarrow b_i} \quad (5.24)$$

onde $reprovacao_{a \rightarrow b_i}$ representa a *reprovação* do agente a em relação a cada um dos agentes b_i , com $b_i \in B$.

$$Y^{aib} = aprovacao_{a_i \rightarrow b} \quad (5.25)$$

onde $aprovacao_{a_i \rightarrow b}$ representa a *aprovação* de cada agente a_i em relação ao agente b , com $a_i \in A$.

$$Y^{a_iB} = \frac{(orgulho_{a_i} + (1 - vergonha_{a_i}))}{2} \quad (5.26)$$

onde $orgulho_{a_i}$ e $vergonha_{a_i}$ representa o *orgulho* e a *vergonha* de cada agente a_i em relação a si mesmo, com $a_i \in A$.

E normalizando esses três valores, temos as três equações dos valores das IEA da RS:

$$IEA^{abi} = \frac{Y^{abi}}{\sum_{b_i \in B} Y^{abi}} \quad (5.27)$$

$$IEA^{aib} = \frac{Y^{aib}}{\sum_{a_i \in A} Y^{aib}} \quad (5.28)$$

$$IEA^{a_iB} = \frac{Y^{a_iB}}{\sum_{a_i \in A} Y^{a_iB}} \quad (5.29)$$

5.3.1 Funcionamento do TrustE

Nessa Seção é apresentada a dinâmica de funcionamento do modelo TrustE através de um cenário de negociação entre agentes.

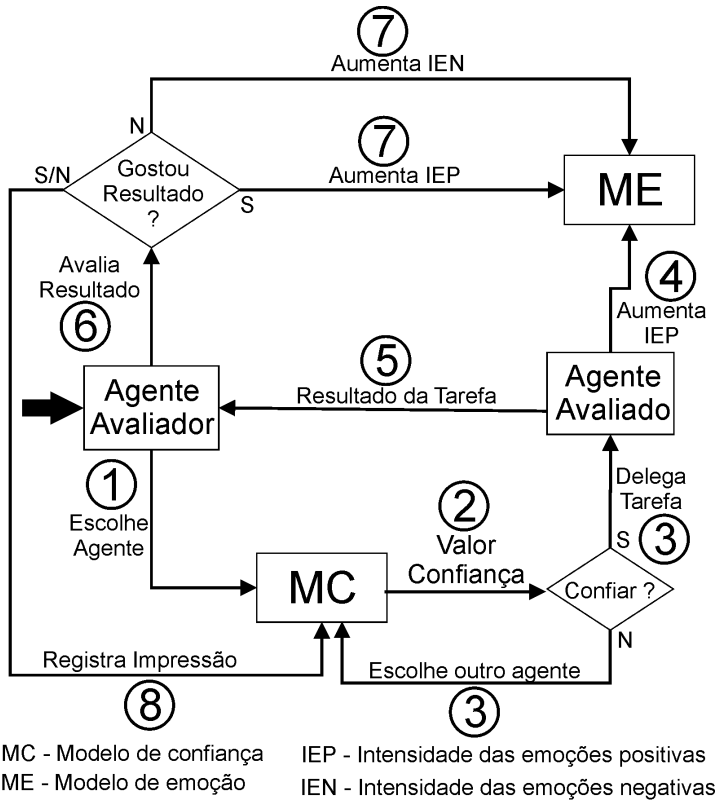
A figura 8 representa uma visão conceitual do TrustE que demonstra o fluxo de ações dos agentes. Para explicar o modelo foi criado um cenário no qual um agente a quer comprar um produto de boa qualidade (A) do agente b , levando em consideração as características presentes na tabela 2. Para simplificar a demonstração, os valores de *limiar* e *potencialidade* são pré-definidos e não mudam com o passar do tempo e o valor de $IP_a(E, t) = 0.5$ para todas as emoções E , com $t \geq 0$.

Tabela 2 – Características dos agentes.

Características	Agente a	Agente b
humor	mal	bom
Potencialidade <i>alegria</i>	0.4	0.7
Potencialidade <i>admiracao</i>	0.4	0.6
Potencialidade <i>reprovacao</i>	0.8	0.4
Potencialidade <i>medo</i>	0.6	0.1
Limiar <i>alegria</i>	0.8	0.3
Limiar <i>admiracao</i>	0.7	0.5
Limiar <i>reprovacao</i>	0.2	0.9
Limiar <i>medo</i>	0.4	0.6
$IMEA_a^E$	0.5	0.5
Produto de boa qualidade	A	A or B

O agente a escolhe um agente vendedor entre todos os vendedores existentes (etapa 1). Para saber se a deve ou não comprar o produto de b , ele irá calcular a confiança em b (etapa 2). Para calcular essa confiança, uma vez que nesse cenário só existem dois agentes, a confiança é calculada somente a partir da RI, através da equação $SR_{a \rightarrow b}(\varphi) = IEE_a^{RI} \cdot R_{a \rightarrow b}(\varphi)$. Supondo que na base de impressões do agente a haja apenas um valor de *impressão*, $W_1 = 0.8$, em relação ao agente b , temos que o valor de $\rho(t, t_1) = 1$. Então $R_{a \rightarrow b}(\varphi) = 1 \cdot (0.8) = 0.8$. Como nesse caso a reputação leva em consideração somente a RI, o valor de $IEE_a^{RI} = 1$, logo $SR_{a \rightarrow b}(\varphi) = 1 \cdot (0.8) = 0.8$. Caso

Figura 8 – Visão Conceitual do Modelo TrustE



esse valor seja considerado baixo, a irá procurar outro agente vendedor, caso contrário, a irá comprar o produto de b (etapa 3). Considerando que 0.8 é um valor alto, o agente a escolhe o agente b como vendedor.

Ao receber a confiança de a e, conseqüentemente ter vendido um produto, o agente b poderá aumentar suas emoções positivas (etapa 4). Para que isso ocorra, primeiramente verifica-se através das equações descritas nas Subseções (5.1.1 e 5.1.2) se b sentiu uma determinada emoção. Supondo que X é um evento tal que $X = \text{vender produto}$ e que b gostaria que X ocorresse, temos que $\text{Satisfeito}_b(X)$ e $\text{Atual}(X)$ são verdadeiros, logo $\text{Alegria}_b(X)$ também é verdadeiro. Assim b ficou feliz por ter vendido um produto e sentiu *alegria*. Como próximo passo b irá calcular a intensidade da *alegria* através da função $I_b(\text{alegria}, 1)$, então $I_b(\text{alegria}, 1) = \max(0, (0.7 - 0.3)) =$

0.4, logo a intensidade da *alegria* = 0.4. Esse valor é utilizado na equação $AI_b(alegria, 1)$ para atualizar a mea_b . Logo $AI_b(alegria, 1) = IMEA_b^{alegria} + \mu \cdot I_b(alegria, 1) = 0.5 + ((0.5) \cdot (0.4)) = 0.7$, e o valor de $IMEA_b^{alegria}$ é atualizado para 0.7.

Passado algum tempo, *b* irá entregar o produto ao agente *a* (etapa 5), e esse por sua vez irá avaliar a qualidade do produto (etapa 6). Como o agente *b* entregou um produto de qualidade B do ponto de vista do agente *a*, e *a* esperava um produto de qualidade A, ele sentirá *reprovacao* em relação a atitude de *b*, supondo que $Reprovacao_a(Y, b) \leftrightarrow Desaprovacao_a(Y) \wedge Acao_b(Y)$ seja verdade, onde $Y = entregar um produto de má qualidade$. Assim $I_a(reprovacao_b, 2) = \max(0, (0.8 - 0.2)) = 0.6$ e $AI_a(reprovacao_b, 2) = IMEA_a^{reprovacao_b} + \mu \cdot I_a(reprovacao_b, 2) = 0.5 + ((0.5) \cdot (0.6)) = 0.8$, logo o valor de $IMEA_a^{reprovacao_b}$ é atualizado para 0.8 (etapa 7). E a avaliação de *a* com respeito a *b* será negativa (etapa 8).

Caso o produto fosse de qualidade A, o agente *a* poderia sentir *admiracao* por *b*, supondo que $Admiracao_a(Y, b)$ seja verdadeiro, onde $Y = entregar um produto de boa qualidade$. Logo a intensidade é calculada através da equação $I_a(admiracao_b, 2) = \max(0, (0.4 - 0.7)) = 0$ e o valor da admiração $AI_a(admiracao_b, 2) = 0.5$, ou seja, apesar do agente *a* ter gostado do produto entregue por *b*, o fato de seu *limiar* da *admiracao* ser maior que a *potencialidade* ($0.7 > 0.4$) o agente *a* não aumenta sua admiração por *b*. Nesse caso a avaliação de *a* em relação a *b* seria positiva (etapa 8).

No início da transação o agente *a* poderia ter sentido *medo* de não receber o produto, supondo que $Insatisfeito_a(X)$ e $Provavel(X)$ sejam verdadeiros, onde $X = não receber o produto$. Assim $Medo_a(X)$ é verdade e a intensidade é dada pela equação $I_a(medo, 1) = \max(0, (0.6 - 0.4)) = 0.2$, e $AI_a(medo, 1) = IMEA_a^{medo} + \mu \cdot I_a(medo, 1) = 0.5 + (0.5 \cdot 0.2) = 0.6$, logo o valor de $IMEA_a^{medo}$ seria atualizado para 0.6.

5.4 AVALIAÇÃO DO MODELO

Para verificar a viabilidade computacional do modelo TrustE, foram simulados e avaliados dois cenários: (1) compra e venda; (2) prestação de serviços. Como o modelo TrustE utiliza oito emoções, para simplificar a simulação, em cada cenário quatro emoções terão o valor da intensidade fixo. No cenário 1 as intensidades das emoções *esperanca*, *medo*, *orgulho* e *vergonha* serão constantes. Já no cenário 2, as emoções que terão intensidades constantes são *alegria*, *angustia*, *aprovacao* e *reprovacao*. As simulações de ambos os cenários foram realizadas com a ferramenta NetLogo (WILENSKY, 1999).

Para facilitar a simulação, ao invés de cada agente guardar uma lista de valores de impressões para cada um dos outros agentes, ele irá armazenar somente um valor para cada agente. Sempre que um agente der uma nova avaliação a respeito de outro, será feita uma média ponderada entre o valor da nova avaliação (peso 0.7) e o valor existente na base de impressões (peso 0.3). O resultado dessa média será o novo valor registrado na base de impressões.

5.4.1 Cenário 1

Esse é um cenário de compra e venda onde existem onze agentes no total, sendo oito agentes compradores e três vendedores. Os agentes compradores pertencem ao grupo *A* enquanto os vendedores ao grupo *B*. O objetivo dos compradores é efetuar o maior número de compras bem sucedidas, enquanto os vendedores buscam efetuar o maior número de vendas.

As intensidades das emoções *esperanca*, *medo*, *orgulho* e *vergonha* terão um valor constante igual a 0.5 para todos os agentes, enquanto as outras quatro emoções assumirão valores de intensidade no intervalo $[0, 1]$. A tabela 3 mostra o perfil emocional de cada agente e as características que dependem de seu humor. Os valores de $IP_a(E, t)$ serão constantes ao longo do tempo t , ao contrário de $H(a, t)$ e $P_a(E, t)$ que terão valores aleatórios dentro do intervalo dado na tabela. Nesse cenário, as emoções são divididas em dois grupos: (i) $E_1 = \{\text{alegria, aprovacao}\}$ e (ii) $E_2 = \{\text{angustia, reprovacao}\}$. O valor de p (fator de atualização) = 0.01 para todas as emoções.

Tabela 3 – Perfil emocional dos agentes

	Humor	$H(a, t)$	$IP_a(E, t)$	$P_a(E, t)$
a_1 e a_2	bom	$[0.7, 1]$	0.8 se $E \in E_1$ 0.3 se $E \in E_2$	$[0.7, 1]$ se $E \in E_1$ $[0, 0.5]$ se $E \in E_2$
a_3, a_4 a_5 e a_6	neutro	$[0.4, 0.7]$	0.5 se $E \in E_1$ 0.5 se $E \in E_2$	$[0.3, 1]$ se $E \in E_1$ $[0.4, 0.8]$ se $E \in E_2$
a_7 e a_8	mal	$[0, 0.4]$	0.4 se $E \in E_1$ 0.9 se $E \in E_2$	$[0, 0.5]$ $E \in E_1$ $[0.5, 1]$ $E \in E_2$
b_1 e b_2	bom	$[0.7, 1]$	0.7 se $E \in E_1$ 0.3 se $E \in E_2$	$[0.7, 1]$ se $E \in E_1$ $[0, 0.5]$ se $E \in E_2$
b_3	mal	$[0, 0.4]$	0.3 se $E \in E_1$ 0.8 se $E \in E_2$	$[0, 0.5]$ se $E \in E_1$ $[0.5, 1]$ $E \in E_2$

A tabela 4 mostra a conduta de cada agente vendedor em relação a

qualidade do produto vendido e do prazo de entrega. Já a tabela 5 mostra os casos em que o agente comprador considera a compra boa (bem sucedida), ruim (mal sucedida) ou neutra (nem boa nem ruim).

Tabela 4 – Conduta dos agentes vendedores

	Qualidade do produto	Entrega
b_1	Ruim	Depois do prazo
b_2	Regular	No prazo
b_3	Boa	Antes do prazo

Tabela 5 – Avaliação da compra

	Vendedor b_1	Vendedor b_2	Vendedor b_3
a_1 e a_2	Neutra	Boa	Boa
a_3, a_4 a_5 e a_6	Ruim	Neutra	Boa
a_7 e a_8	Ruim	Ruim	Neutra

Levando em consideração os valores da tabela 5, sempre que um agente comprador avaliar a compra como ruim, ele irá sentir emoções negativas (angústia e reprovação), quando a compra for considerada boa, o agente irá sentir emoções positivas (alegria e aprovação). E quando for neutra, o agente irá sentir emoções positivas ou negativas. Por sua vez, um vendedor irá sentir alegria quando realizar uma venda e angústia quando deixar de vender por ser considerado pouco confiável.

A tabela 6 mostra o perfil da avaliação (valor da impressão - W_i) dos agentes compradores em relação aos agentes vendedores após a realização de uma compra. Esses valores levam em consideração a avaliação da compra presente na tabela 5. O agente comprador dará uma avaliação maior ao vendedor quando a compra for boa e uma avaliação menor quando for ruim.

Tabela 6 – Valor das avaliações (impressões) dos compradores

	Vendedor b_1	Vendedor b_2	Vendedor b_3
a_1 e a_2	$[-0.5, 0]$	$[0, 0.5]$	$[0.5, 1]$
a_3, a_4 a_5 e a_6	$[-1, 0]$	$[0, 0.4]$	$[0.4, 0.9]$
a_7 e a_8	$[-1, -0.5]$	$[-0.5, 0.3]$	$[0.3, 0.7]$

Durante a simulação, os agentes seguem o ciclo de interação definido

a seguir:

1. Um agente *a* (vendedor) escolhe, de forma aleatória, um dos compradores (agente *b*) para efetuar uma compra.
2. *a* calcula a confiança em *b*. Se a confiança for satisfatória *a* irá comprar o produto de *b*, e esse aumentará a sua *alegria* por ter vendido um produto. Caso a confiança seja insuficiente, *a* escolherá outro vendedor e *b* sentirá *angústia* por não ter realizado a venda. Essa informação poderá ser utilizada futuramente pelo agente, nas suas interações.
3. Caso *a* não encontre nenhum outro vendedor confiável ele irá parar de comprar.
4. Passado algum tempo, *a* recebe o produto de *b*. Se *a* considerar a compra boa, ele aumentará sua *alegria* e sua *aprovação* em relação a *b*. Caso *a* considere a compra ruim, ele irá aumentar sua *angústia* e sua *reprovação* em relação a *b*. E caso *a* considere a compra neutra ele sentirá, de forma aleatória, emoções positivas ou negativas.
5. O agente *a* avalia *b*, e grava sua avaliação (W_i) na base de impressões.
6. O agente *a* aguarda um intervalo de tempo *t* para realizar a próxima compra.

Para obter-se uma análise mais ampla desse cenário, foram definidos três parâmetros: (i) *max-compra* (número máximo de compras que um comprador pode realizar), (ii) *valor-inicial-impressoes* (o valor inicial das impressões de que cada agente comprador em relação aos vendedores) e (iii) *conf-min-para-comprar* (o valor mínimo da confiança que o comprador deve ter em relação ao vendedor para efetuar a compra). O tempo que cada comprador leva para realizar uma nova compra varia de $5t$ a $10t$.

Foram realizadas três simulações desse cenário, sendo que em cada simulação foram utilizados conjuntos de valores distintos para os três parâmetros. Ao final de cada simulação, foram avaliadas as emoções *alegria*, *angustia*, *aprovacao* e *reprovacao* dos agentes compradores, e as emoções *alegria* e *angustia* dos vendedores. Também foi analisado o total de vendas de cada vendedor e as avaliações dos compradores em relação aos vendedores.

A tabela 7 apresenta os valores dos parâmetros em cada uma das três simulações.

A seguir são apresentados os resultados da Simulação 1 desse Cenário. As simulações 2 e 3 estão descritas no Apêndice A.

Tabela 7 – Configuração dos parâmetros

Parâmetros	Simulação 1	Simulação 2	Simulação 3
<i>max-compra</i>	10	20	30
<i>valor-inicial-impressoes</i>	1	0.5	0
<i>conf-min-para-comprar</i>	0.5	0	0

5.4.1.1 Simulação 1

Tabela 8 – Intensidade das emoções dos agentes

	<i>ale</i>	<i>ang</i>	<i>apr_{b1}</i>	<i>apr_{b2}</i>	<i>apr_{b3}</i>	<i>rep_{b1}</i>	<i>rep_{b2}</i>	<i>rep_{b3}</i>
<i>a₁</i>	0.98	0.3	0.8	0.8	0.98	0.3	0.3	0.3
<i>a₂</i>	0.98	0.3	0.8	0.8	0.98	0.3	0.3	0.3
<i>a₃</i>	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
<i>a₄</i>	0.69	0.5	0.5	0.69	0.5	0.5	0.5	0.5
<i>a₅</i>	0.71	0.5	0.5	0.5	0.71	0.5	0.5	0.5
<i>a₆</i>	0.92	0.5	0.5	0.5	0.92	0.5	0.5	0.5
<i>a₇</i>	0.4	0.9	0.4	0.4	0.4	0.9	0.9	0.9
<i>a₈</i>	0.4	0.9	0.4	0.4	0.4	0.9	0.9	0.9
<i>b₁</i>	0.7	0.3	x	x	x	x	x	x
<i>b₂</i>	0.7	0.3	x	x	x	x	x	x
<i>b₃</i>	0.3	0.8	x	x	x	x	x	x

ale : alegria, *ang* : angustia, *apr* : aprovacao, *rep* : reprovacao

Analizando a tabela 8 referente as memórias de emoções dos agentes, observa-se que o resultado esperado foi alcançado. Os agentes de bom humor terminaram a simulação com valores altos para emoções positivas e baixo para negativas, enquanto os agentes de mau humor tiveram a intensidade de suas emoções positivas baixas e as intensidades das emoções negativas altas.

0

Tabela 9 – Total de compras

	<i>a₁</i>	<i>a₂</i>	<i>a₃</i>	<i>a₄</i>	<i>a₅</i>	<i>a₆</i>	<i>a₇</i>	<i>a₈</i>
Bem Sucedida	9	9	6	7	8	8	0	0
Mal Sucedida	0	0	1	1	1	1	2	2
Neutra	1	1	1	3	1	1	6	5

A tabela 9 demonstra que os agentes, de bom humor e de humor neutro, conseguiram atingir o seu objetivo que era realizar o maior número de compras bem sucedidas. Os agentes de mau humor não obtiveram nenhuma compra bem sucedida pois para esses agentes nenhum dos vendedores realiza uma venda considerada boa (ver tabela 5).

A tabela 10 mostra o total de vendas de cada vendedor. Como era esperado, o vendedor b_3 teve um número de vendas bem expressivo. Já os vendedores b_1 e b_2 , por terem uma qualidade de venda inferior, obtiveram um número baixo nas vendas.

Tabela 10 – Total de vendas

	b_1	b_2	b_3
Total de vendas	8	9	56

Tabela 11 – Base de impressões dos compradores em relação aos vendedores

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
b_1	0.02	0.02	0.16	-0.26	-0.33	0.23	-0.05	-0.33
b_2	0.3	0.44	0.37	0.17	0.37	0.3	0.3	0.3
b_3	0.87	0.87	0.48	0.8	0.72	0.51	0.4	0.44

As impressões presentes na base de impressões de cada comprador ao término da simulação (tabela 11), demonstram que o vendedor b_1 , que é o “pior vendedor” tem valores baixos de avaliação, enquanto os outros dois tem valores médios e altos.

5.4.2 Cenário 2

Esse é um cenário de prestação de serviço, onde agentes médicos prestam serviços aos agentes pacientes. No grupo A , existem 10 agentes pacientes, enquanto no grupo B , cinco agentes médicos. O objetivo dos pacientes é conseguir realizar o maior número de consultas bem sucedidas. Enquanto o objetivo dos médicos é realizar o maior número de consultas.

As intensidades das emoções *alegria*, *angustia*, *aprovação* e *reprovação* terão um valor constante = 0.5 para todos os agentes, enquanto as outras quatro emoções terão valores de intensidade no intervalo $[0, 1]$. A tabela 12 mostra o perfil emocional de cada agente e as características que dependem de seu humor. Assim como no cenário 1, os valores de $IP_a(E, t)$

serão constantes ao longo do tempo t , ao contrário de $H(a,t)$ e $P_a(E,t)$ que terão valores aleatórios dentro do intervalo dado na tabela.

Tabela 12 – Perfil emocional dos agentes

	Humor	$H(a,t)$	$IP_a(E,t)$	$P_a(E,t)$
a_1 e a_2 a_3	bom	$[0.8, 1]$	0.7 se $E \in E_1$ 0.4 se $E \in E_2$	$[0.8, 1]$ se $E \in E_1$ $[0, 0.5]$ se $E \in E_2$
$a_4, a_5,$ a_6 e a_7	neutro	$[0, 1]$	0.5 se $E \in E_1$ 0.5 se $E \in E_2$	$[0, 1]$ se $E \in E_1$ $[0, 1]$ se $E \in E_2$
a_8 e a_9 a_{10}	mal	$[0, 0.5]$	0.2 se $E \in E_1$ 0.8 se $E \in E_2$	$[0, 0.4]$ $E \in E_1$ $[0.5, 0.8]$ $E \in E_2$
b_1, b_2	bom	$[0.7, 1]$	0.7 se $E \in E_1$ 0.4 se $E \in E_2$	$[0.8, 1]$ se $E \in E_1$ $[0, 0.5]$ se $E \in E_2$
b_3	neutro	$[0, 1]$	0.5 se $E \in E_1$ 0.5 se $E \in E_2$	$[0, 1]$ se $E \in E_1$ $[0, 1]$ se $E \in E_2$
b_4 e b_5	mal	$[0, 0.5]$	0.2 se $E \in E_1$ 0.8 se $E \in E_2$	$[0, 0.4]$ se $E \in E_1$ $[0.5, 0.8]$ $E \in E_2$

Nesse cenário $E_1 = \{\text{esperanca, orgulho}\}$, $E_2 = \{\text{vergonha, medo}\}$ e p (fator de atualização) = 0.01 para todas as emoções.

A tabela 13 mostra a conduta de cada médico em relação a qualidade da prestação de serviço (qualidade e preço da consulta). Enquanto a tabela 14 mostra os casos em que o agente paciente considera a consulta boa (bem sucedida), ruim (mal sucedida) ou neutra (nem boa nem ruim).

Tabela 13 – Conduta dos agentes médicos

	Qualidade da Consulta	Preço da Consulta
b_1 e b_2	Boa	Alto
b_3 e b_4	Regular	Baixo
b_5	Ruim	Médio

Levando em consideração os valores da tabela 14, sempre que um agente paciente avaliar uma consulta como ruim, ele irá sentir emoções negativas (vergonha e medo). Quando for considerada boa, o agente irá sentir emoções positivas (orgulho e esperança). E quando for neutra, o agente irá sentir emoções positivas ou negativas de forma aleatória. Os agentes médicos sentirão orgulho quando realizaram uma consulta, e vergonha, quando deixarem de atender um paciente por não serem considerados confiáveis.

A tabela 15 mostra o perfil da avaliação (valor da impressão - W_i) dos agentes pacientes em relação aos agentes médicos.

Tabela 14 – Avaliação da consulta

	Méd b_1	Méd b_2	Méd b_3	Méd b_4	Méd b_5
a_1, a_2 e a_3	Boa	Neutra	Boa	Neutra	Ruim
$a_4, a_5,$ a_6 e a_7	Neutra	Boa	Boa	Ruim	Neutra
a_8, a_9 e a_{10}	Ruim	Boa	Neutra	Neutra	Ruim

Tabela 15 – Valor das avaliações (impressões)

	Med b_1	Med b_2	Med b_3	Med b_4	Med b_5
a_1, a_2 e a_3	[0.5, 1]	[0, 0.8]	[0.3, 1]	[0, 0.8]	[-1, -0.5]
$a_4, a_5,$ $a_6, a_7,$	[-0.2, 0.6]	[0, 1]	[0.4, 1]	[-1, 0]	[-0.5, 0.5]
a_8, a_9 e a_{10}	[-0.8, 0]	[0.6, 1]	[0, 0.4]	[-0.3, 0.4]	[-0.5, 0]

Durante a simulação os agentes seguem o ciclo de interação definido a seguir:

1. Um agente a (paciente) escolhe um dos agentes pertencentes ao grupo B (médicos) para marcar uma consulta aleatoriamente.
2. a calcula a confiança em b , se a confiança for satisfatória a irá marcar a consulta com b , caso contrário, a escolherá outro médico, e esse por sua vez, irá sentir *vergonha* por não ter sido considerado confiável.
3. Caso a não encontre nenhum outro médico confiável ele irá parar de marcar consultas.
4. Após o médico realizar a consulta, ele irá aumentar o seu *orgulho* por ter atendido um paciente, e poderá usar esta informação, no futuro para, por exemplo, aumentar o preço da sua consulta.
5. Se a considerar a consulta boa, ele aumentará sua *esperança* e seu *orgulho*. Caso a considere a consulta ruim, ele irá aumentar seu *medo* e sua *vergonha*. E caso a considere a consulta neutra, ele aumentará suas emoções positivas ou negativas.
6. O agente a avalia b , e grava sua avaliação (W_i) na base de impressões.

7. O agente a aguarda um intervalo de tempo t para marcar a próxima consulta.

Assim como no cenário 1, nesse cenário foram definidos três parâmetros: (i) *max-consultas*, (ii) *valor-inicial-impressoes* e (iii) *confianca-min-para-consultar*. O tempo que cada paciente leva para marcar uma nova consulta varia de $10t$ a $15t$.

Nesse cenário também foram realizadas três simulações, sendo que em cada simulação foram utilizados conjuntos de valores distintos para os três parâmetros. Ao final das três simulações foram avaliadas as emoções *esperança*, *medo*, *orgulho* e *vergonha* dos agentes pacientes, e as emoções *orgulho* e *vergonha* dos médicos. Também foi analisado o total de consultas realizadas por cada médico as avaliações dos pacientes em relação aos médicos.

A tabela 16 apresenta os valores dos parâmetros em cada simulação.

Tabela 16 – Configuração dos parâmetros

Parâmetros	Simulação 1	Simulação 2	Simulação 3
<i>max-consultas</i>	5	10	15
<i>valor-inicial-impressoes</i>	0.8	0.5	0.2
<i>confianca-min-para-consultar</i>	0.7	0.3	0

A seguir são apresentados os resultados da Simulação 1 desse Cenário. As simulações 2 e 3 estão descritas no Apêndice B.

5.4.2.1 Simulação 1

Analisando a tabela 17 referentes às memórias de emoções dos agentes, observa-se que o resultado esperado foi alcançado. Os agentes de bom humor terminaram a simulação com valores altos para emoções positivas e baixo para negativas, enquanto os agentes de mau humor tiveram a intensidade de suas emoções positivas baixas e as intensidades das emoções negativas altas.

A tabela 18 demonstra que os agentes, de bom humor e de humor neutro, conseguiram atingir o seu objetivo que era realizar o maior número de consultas bem sucedidas. A tabela 19 mostra o total de consultas de cada médico. O médico b_2 foi o que teve maior número de consultas realizadas por ser um dos melhores médicos.

Tabela 17 – Intensidade das emoções dos agentes

	<i>esperanca</i>	<i>medo</i>	<i>orgulho</i>	<i>vergonha</i>
a_1	0.98	0.4	1	0.4
a_2	0.95	0.4	0.94	0.4
a_3	0.89	0.4	0.88	0.4
a_4	0.53	0.5	0.5	0.5
a_5	0.5	0.5	0.65	0.5
a_6	0.5	0.5	0.61	0.5
a_7	0.5	0.5	0.5	0.5
a_8	0.2	0.84	0.2	0.9
a_9	0.2	0.8	0.2	0.8
a_{10}	0.2	0.8	0.2	0.82
b_1	x	x	0.89	0.4
b_2	x	x	0.9	0.4
b_3	x	x	0.83	0.5
b_4	x	x	0.2	0.88
b_5	x	x	0.2	0.9

Tabela 18 – Total de consultas dos pacientes

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Bem Sucendida	2	3	2	2	2	2	2	3	1	2
Mal Sucendida	1	0	1	1	1	1	1	2	1	2
Neutra	2	2	2	1	2	2	1	0	2	1

Tabela 19 – Total de consultas dos médicos

	b_1	b_2	b_3	b_4	b_5
Total de consultas	10	13	10	8	6

As impressões presentes na base de impressões de cada paciente ao término da simulação (tabela 20), demonstra que o valor das impressões dos médicos b_1 , b_3 , b_4 e b_5 mantiveram-se praticamente iguais aos valores iniciais, uma vez que, eles realizaram poucas consultas.

Fazendo uma avaliação geral das simulações do Cenário 1 e 2, observa-se que os vendedores e médicos ruins são menos procurados à medida que novas interações são realizadas. Isso ocorre porque as avaliações recebidas por esses agentes, normalmente, são ruins. Em relação as impressões registradas pelos agentes, nota-se que aqueles que possuem uma

Tabela 20 – Base de impressões dos pacientes em relação aos médicos

	b_1	b_2	b_3	b_4	b_5
a_1	0.94	0.31	0.45	0.24	-0.25
a_2	0.66	0.52	0.52	0.73	0.8
a_3	0.59	0.73	0.52	0.31	-0.11
a_4	0.38	0.45	0.73	0.17	0.8
a_5	0.24	0.73	0.59	0.24	0.38
a_6	0.52	0.8	0.66	0.3	0.1
a_7	0.17	0.38	0.59	-0.25	0.8
a_8	0.3	0.71	0.8	0.8	0.03
a_9	0.3	0.73	0.45	0.17	0.8
a_{10}	-0.18	0.68	0.24	0.8	0.24

boa conduta tendem a receber uma avaliação mais alta.

6 CONSIDERAÇÕES FINAIS

Esse trabalho apresentou o modelo híbrido TrustE, que é um modelo de confiança, baseado no Regret, que utiliza as emoções do agente no mecanismo de cálculo da confiança, visando capturar a complexidade do raciocínio humano e flexibilizar as avaliações quantitativas, introduzindo elementos de natureza qualitativa ao modelo. As emoções vivenciadas pelos agentes podem ser originárias do próprio modelo TrustE como de sensações externas ao modelo. Para viabilizar a inserção das emoções internas ao modelo TrustE, foi necessário a formalização de oito tipos de emoções presentes no modelo OCC, a modelagem de funções que representam a intensidade dessas emoções e a definição de funções que gerenciam a memória de emoções dos agentes.

Como as emoções estão fortemente relacionadas com o ato de confiar e os modelos de confiança presentes na literatura adotam uma abordagem estritamente quantitativa, a proposta de um modelo de confiança baseado em emoções surge como uma alternativa, a fim de tornar o ato de confiar mais subjetivo e próximo do comportamento humano.

Analizando o modelo proposto e as simulações realizadas, verifica-se que a inserção de emoções no cálculo de confiança é uma alternativa eficaz comparada aos modelos que não utilizam emoções, além de tornar o modelo mais abrangente. Um agente sentindo alegria, pode, por exemplo, não levar em consideração as impressões obtidas de outros agentes mesmo quando essas impressões forem melhores que as suas.

O modelo TrustE utiliza as avaliações dos agentes em conjunto com suas emoções para tornar o processo de tomada de decisão mais subjetivo, entretanto a utilização de emoções torna a implementação do modelo mais complexa se comparada a um modelo de confiança que não utilize emoções.

O objetivo de propor um modelo de confiança baseado em emoções, que utiliza informações de natureza simbólica foi alcançado com sucesso. Para alcançar tal objetivo, foi realizado um estudo dos modelos de confiança em agentes, do modelo OCC e de modelos de emoções para agentes. Esse estudo permitiu a formalização das emoções e das equações de intensidade das emoções utilizadas no modelo proposto. Entre as dificuldades encontradas para alcançar tal objetivo, podemos citar: (i) decidir quais emoções devem ser levadas em consideração na formação dos pesos em cada uma das funções do cálculo de confiança; (ii) como modelar as funções relativas à intensidade das emoções e as funções que gerenciam a memória de emoções dos agentes.

O modelo TrustE só utiliza uma função para calcular a intensidade de todas as emoções do modelo. Esse com certeza é um ponto que precisa

ser melhorado em trabalhos futuros, para que cada emoção tenha uma função específica de intensidade. Outro ponto fraco, é o fato da escolha das emoções, utilizadas nas equações que formam os pesos *IEE* e *IEA*, ter sido feita sem um estudo mais aprofundado, que possibilitaria uma escolha mais adequada dessas emoções, tornando o modelo mais preciso.

Por fim, acredita-se que a inserção de emoções e a utilização de sua intensidade possa propiciar um maior realismo aos modelos de confiança, pois a tomada de decisão do agente estará diretamente ligada a seu estado emocional. Devido a flexibilidade presente no TrustE, ele pode ser expandido para comportar um maior número de emoções, sendo necessário apenas alguns ajustes no modelo.

6.1 TRABALHOS FUTUROS

Como trabalhos futuros sugere-se:

- estudar com maior profundidade como as emoções impactam o comportamento dos agentes;
- aumentar o número de emoções do modelo, e formalizá-las de forma mais precisa, propiciando um maior realismo;
- remodelar as funções relacionadas as emoções e a memória de emoções a fim de tornar o modelo mais complexo e próximo do comportamento humano;
- avaliar a tomada de decisão de confiar ou não dos agentes, considerando a influência das emoções externas ao modelo TrustE;
- desenvolver um arcabouço (framework) que implemente o modelo e que possa ser facilmente incorporado em sistemas multiagente.

REFERÊNCIAS

ADAM, C. et al. Occ's emotions: a formalization in a bdi logic. **Artificial Intelligence: Methodology, Systems, and Applications: 12th International Conference**, Springer, p. 24–32, 2006.

DASTANI, M.; MEYER, J.-J. C. Programming agents with emotions. In: **ECAI 2006: 17th European Conference on Artificial Intelligence**. [S.l.: s.n.].

DONG-HUYNHA, T.; JENNINGS, N.; SHADBOLT, N. Fire: An integrated trust and reputation model for open multi-agent systems. In: IOS PR INC. **ECAI 2004: 16th European Conference on Artificial Intelligence, August 22-27**. [S.l.], 2004. v. 110, p. 18.

FERBER, J. **Multi-agent systems: an introduction to distributed artificial intelligence**. [S.l.]: Addison-Wesley Longman Publishing Co., Inc., 1999.

FOSTER, I.; KESSELMAN, C. **The grid: blueprint for a new computing infrastructure**. [S.l.]: Morgan Kaufmann, 2004.

GUEMKAM, G. et al. Metamodel for reputation based agents system: Case study for electrical distribution scada design. In: **Proceedings of the 6th International Conference on Security of Information and Networks**. New York, NY, USA: ACM, 2013. (SIN '13), p. 251–255. ISBN 978-1-4503-2498-4. Disponível em: <http://doi.acm.org/10.1145/2523514.2523543>.

HERZIG, A.; LONGIN, D. Cl intention revisited. In: **KR**. [S.l.: s.n.], 2004. p. 527–535.

JIANG, H.; VIDAL, J. M.; HUHNS, M. N. Ebdi: An architecture for emotional agents. In: **Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems**. New York, NY, USA: ACM, 2007. (AAMAS '07), p. 11:1–11:3. ISBN 978-81-904262-7-5. Disponível em: <http://doi.acm.org/10.1145/1329125.1329139>.

LÓPEZ, M. R. et al. **Definición de un modelo unificador del concepto de confianza mediante esquemas preconceptuales**. Tese (Doutorado) — Universidad Nacional de Colombia, Sede Medellín, 2011.

LORINI, E. Agents with emotions: a logical perspective. **ALP Newsletter**, s. n., v. 12, n. 2-3, 2008.

LU, G. et al. A review on computational trust models for multi-agent systems. **The Open Information Science Journal**, Bentham Open, 2009.

MARSH, S. **Formalising Trust as a Computational Concept**. Tese (Doutorado) — Department of Mathematics and Computer Science, University of Stirling, 1994.

MEYER, J.-J. C. Reasoning about emotional agents. **International journal of intelligent systems**, Wiley Online Library, v. 21, n. 6, p. 601–619, 2006.

MEYER, J.-J. C.; HOEK, W. van der; LINDER, B. van. A logical approach to the dynamics of commitments. **Artificial Intelligence**, Elsevier, v. 113, n. 1, p. 1–40, 1999.

MICHAELIS - Moderno Dicionário da Língua Portuguesa. São Paulo: [s.n.], 2012. ISBN 978-85-06-06953-0.

ORTONY, A.; CLORE, G.; COLLINS, A. **The cognitive structure of emotions**. [S.l.]: Cambridge university press, 1990.

OSTROM, E. A behavioral approach to the rational choice theory of collective action. **The American Political Science Review**, s. n., v. 92, n. 1, p. 1–22, Mar 1998.

RUSSELL, S.; NORVIG, P. **Inteligência Artificial: tradução da segunda edição**. [S.l.]: Elsevier: Campus, 2004.

SABATER, J.; SIERRA, C. Regret: Reputation in gregarious societies. In: **Proceedings of the Fifth International Conference on Autonomous Agents**. New York, NY, USA: ACM, 2001. (AGENTS '01), p. 194–195. ISBN 1-58113-326-X. Disponível em: <<http://doi.acm.org/10.1145/375735.376110>>.

SABATER, J.; SIERRA, C. Review on computational trust and reputation models. **Artificial Intelligence Review**, Kluwer Academic Publishers, Norwell, MA, USA, v. 24, n. 1, p. 33–60, 2005. ISSN 0269-2821. Disponível em: <<http://dx.doi.org/10.1007/s10462-004-0041-5>>.

SABATER-MIR, J.; PAOLUCCI, M. On representation and aggregation of social evaluations in computational trust and reputation models. **International Journal of Approximate Reasoning**, v. 46, n. 3, p. 458 – 483, 2007. ISSN 0888-613X. Special Section: Aggregation Operators.

Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0888613X07000059>>.

SENSOY, M. et al. Reasoning about uncertain information and conflict resolution through trust revision. In: **Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems**. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2013. (AAMAS '13), p. 837–844. ISBN 978-1-4503-1993-5. Disponível em: <<http://dl.acm.org/citation.cfm?id=2484920.2485053>>.

STEUNEBRINK, B.; DASTANI, M.; MEYER, J. The occ model revisited. In: **Proceedings of the 4th Workshop on Emotion and Computing**. [S.l.]: s. n., 2009. v. 65, p. 2047–2056.

STEUNEBRINK, B. R.; DASTANI, M.; MEYER, J. A logic of emotions for intelligent agents. In: AAAI PRESS. **Proceedings of the National Conference on Artificial Intelligence**. [S.l.], 2007. v. 22, n. 1, p. 142.

STEUNEBRINK, B. R.; DASTANI, M.; MEYER, J.-J. C. Towards a quantitative model of emotions for intelligent agents. In: **Proceedings of the 2nd Workshop on Emotion and Computing-Current Research and Future Impact**. [S.l.]: s. n., 2007.

STEUNEBRINK, B. R.; DASTANI, M.; MEYER, J.-J. C. A formal model of emotion triggers: an approach for bdi agents. **Synthese**, Springer, v. 185, n. 1, p. 83–129, 2012.

TEACY, W. T. L. et al. Travos: Trust and reputation in the context of inaccurate information sources. **Autonomous Agents and Multi-Agent Systems**, Springer, v. 12, n. 2, p. 183–198, 2006.

TURRINI, P.; MEYER, J.-J.; CASTELFRANCHI, C. Rational agents that blush. **Affective Computing and Intelligent Interaction**, Springer, p. 314–325, 2007.

WILENSKY, U. **NetLogo**. 1999. Disponível em: <<http://ccl.northwestern.edu/netlogo>>.

WOOLDRIDGE, M. **An introduction to multiagent systems**. [S.l.]: Wiley, 2009.

YU, B.; SINGH, M. Searching social networks. In: ACM. **Proceedings of the second international joint conference on Autonomous agents and multiagent systems**. [S.l.], 2003. p. 65–72.

ZACHARIA, G.; MOUKAS, A.; MAES, P. Collaborative reputation mechanisms for electronic marketplaces. **Decision Support Systems**, Elsevier, v. 29, n. 4, p. 371–388, 2000.

APÊNDICE A – Simulações Cenário 1

A.1 SIMULAÇÃO 2

Tabela 21 – Intensidade das emoções dos agentes

	<i>ale</i>	<i>ang</i>	<i>apr_{b₁}</i>	<i>apr_{b₂}</i>	<i>apr_{b₃}</i>	<i>rep_{b₁}</i>	<i>rep_{b₂}</i>	<i>rep_{b₃}</i>
<i>a₁</i>	0.99	0.3	0.8	0.99	0.82	0.3	0.3	0.3
<i>a₂</i>	0.99	0.3	0.8	0.91	0.99	0.3	0.3	0.3
<i>a₃</i>	0.69	0.5	0.5	0.53	0.5	0.5	0.5	0.5
<i>a₄</i>	0.56	0.5	0.5	0.5	0.5	0.5	0.5	0.5
<i>a₅</i>	0.7	0.5	0.5	0.63	0.5	0.5	0.5	0.5
<i>a₆</i>	0.64	0.5	0.5	0.64	0.5	0.5	0.5	0.5
<i>a₇</i>	0.4	0.9	0.4	0.4	0.4	0.9	0.9	0.9
<i>a₈</i>	0.4	0.94	0.4	0.4	0.4	0.9	0.9	0.94
<i>b₁</i>	0.7	0.3	x	x	x	x	x	x
<i>b₂</i>	0.97	0.3	x	x	x	x	x	x
<i>b₃</i>	0.3	0.8	x	x	x	x	x	x

ale : alegria, *ang* : angustia, *apr* : aprovacao, *rep* : reprovacao

Analisando a tabela 21 referentes as memórias de emoções dos agentes, assim como na Simulação 1, observa-se que os agentes de bom humor terminaram a simulação com valores altos para emoções positivas e baixo para negativas, enquanto os agentes de humor neutro ficaram com valores médios tanto para emoções positivas como negativas.

Tabela 22 – Total de compras

	<i>a₁</i>	<i>a₂</i>	<i>a₃</i>	<i>a₄</i>	<i>a₅</i>	<i>a₆</i>	<i>a₇</i>	<i>a₈</i>
Bem Sucedida	18	19	10	7	11	10	0	0
Mal Sucedida	0	0	1	1	1	1	3	5
Neutra	2	1	9	12	8	9	17	15

Analisando a tabela 22, verifica-se que os agentes de bom humor alcançaram um alto número de compras bem sucedidas, como era esperado.

Conforme pode ser observados na tabela 23, o vendedor *b₁* é o que apresenta o menor número de vendas, visto que sua conduta é a pior entre os vendedores. Por sua vez, o vendedor *b₃* possui o maior número de vendas, já que possui a melhor conduta. Da mesma maneira que na Simulação 1, o vendedor *b₃* tem os maiores valores de impressões (tabela 24).

Tabela 23 – Total de vendas

	b_1	b_2	b_3
Total de vendas	9	59	92

Tabela 24 – Base de impressões dos compradores em relação aos vendedores

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
b_1	-0.23	-0.2	-0.2	-0.27	-0.48	-0.2	-0.27	-0.55
b_2	0.25	0.35	0.01	0.1	0.34	0.32	-0.24	-0.29
b_3	0.83	0.94	0.57	0.72	0.78	0.75	0.49	0.58

A.2 SIMULAÇÃO 3

Tabela 25 – Intensidade das emoções dos agentes

	ale	ang	apr_{b_1}	apr_{b_2}	apr_{b_3}	rep_{b_1}	rep_{b_2}	rep_{b_3}
a_1	0.94	0.3	0.8	0.86	0.92	0.3	0.3	0.3
a_2	0.95	0.3	0.8	0.9	0.88	0.3	0.3	0.3
a_3	0.82	0.5	0.5	0.8	0.5	0.5	0.5	0.5
a_4	0.6	0.66	0.5	0.6	0.5	0.5	0.66	0.5
a_5	0.56	0.5	0.5	0.55	0.51	0.5	0.5	0.5
a_6	0.4	0.9	0.4	0.4	0.4	0.9	0.9	0.9
a_7	0.82	0.5	0.5	0.5	0.8	0.5	0.5	0.5
a_8	0.4	0.98	0.4	0.4	0.4	0.9	0.9	0.98
b_1	0.79	0.3	x	x	x	x	x	x
b_2	0.95	0.3	x	x	x	x	x	x
b_3	0.3	0.8	x	x	x	x	x	x

ale : alegria, ang : angustia, apr : aprovacao, rep : reprovacao

Nessa simulação a memória de emoções dos agentes tem a mesma característica das duas simulações anteriores.

A tabela 27 mostra o total de vendas de cada vendedor. Como era esperado, o vendedor b_3 teve o maior número de vendas, seguido pelo vendedor b_2 . O vendedor b_1 teve um número de vendas bem baixo, visto que ele tem uma má conduta em relação as vendas.

As impressões presentes na base de impressões de cada comprador ao término da simulação, demonstra que o vendedor b_1 teve valores baixos de

Tabela 26 – Total de compras

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
Bem Sucedida	30	29	15	13	16	13	0	0
Mal Sucedida	0	0	1	1	0	1	4	3
Neutra	0	1	14	16	14	16	26	27

Tabela 27 – Total de vendas

	b_1	b_2	b_3
Total de vendas	6	89	145

Tabela 28 – Base de impressões dos compradores em relação aos vendedores

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
b_1	0	-0.28	-0.56	-0.07	0	-0.28	-0.7	-0.49
b_2	0.35	0.26	0.21	0.07	0.15	0.1	-0.24	-0.29
b_3	0.6	0.66	0.63	0.62	0.7	0.42	0.45	0.54

avaliação, uma vez que, b_1 não prestou um bom serviço.

Analisando as três simulações desse cenário, observa-se que a medida que o número de compras aumenta, a quantidade de vendas realizadas pelos vendedores b_2 e b_3 aumenta, enquanto a quantidade de vendas do vendedor b_1 praticamente não muda. Isso ocorre porque com um número maior de interações, as avaliações a respeito dos vendedores tornam-se mais precisas, fazendo com que os compradores procurem somente os vendedores que possuam uma boa conduta.

APÊNDICE B – Simulações Cenário 2

B.1 SIMULAÇÃO 2

Tabela 29 – Intensidade das emoções dos agentes

	<i>esperanca</i>	<i>medo</i>	<i>orgulho</i>	<i>vergonha</i>
a_1	0.83	0.4	0.79	0.4
a_2	0.94	0.4	0.95	0.4
a_3	0.97	0.4	0.95	0.4
a_4	0.96	0.5	0.7	0.5
a_5	0.5	0.5	0.5	0.5
a_6	0.75	0.5	0.62	0.5
a_7	0.75	0.5	0.95	0.5
a_8	0.2	0.8	0.2	0.8
a_9	0.2	0.8	0.2	0.8
a_{10}	0.2	0.84	0.2	0.84
b_1	x	x	0.79	0.4
b_2	x	x	0.89	0.4
b_3	x	x	0.83	0.5
b_4	x	x	0.2	0.83
b_5	x	x	0.2	0.88

Analisando a tabela 29 referentes as memórias de emoções dos agentes, observa-se que o resultado esperado foi alcançado.

Tabela 30 – Total de consultas dos pacientes

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Bem Sucedida	4	6	3	7	6	6	4	5	2	2
Mal Sucedida	1	1	1	1	1	1	1	2	2	2
Neutra	5	3	6	2	3	3	5	3	6	6

A tabela 30 mostra o total de consultas de cada paciente. Como esperado, a maioria das consultas realizadas pelos pacientes foram bem sucedidas.

Tabela 31 – Total de consultas por médico

	b_1	b_2	b_3	b_4	b_5
Total de consultas	15	21	37	16	11

Tabela 32 – Base de impressões dos pacientes em relação aos médicos

	b_1	b_2	b_3	b_4	b_5
a_1	0.85	0.57	0.84	0.55	-0.2
a_2	0.72	0.55	0.76	0.09	-0.55
a_3	0.71	0.61	0.58	0.21	-0.48
a_4	0.22	0.89	0.71	0.15	0.43
a_5	0.38	0.43	0.52	-0.48	0.08
a_6	0.36	0.22	0.53	-0.48	0.08
a_7	0.02	0.15	0.89	-0.34	-0.15
a_8	-0.27	0.65	0.13	0.15	-0.06
a_9	-0.06	0.66	0.11	0.34	-0.06
a_{10}	-0.02	0.7	0.21	-0.06	0.015

Diferentemente da primeira simulação, as impressões presentes na base de impressões de cada paciente ao término dessa simulação (tabela 32), está bem diferente do valor inicial. Isso ocorreu devido ao fato do número de consultas por pacientes ser maior, aliado ao aumento no número de consultas realizadas por médico.

B.2 SIMULAÇÃO 3

Tabela 33 – Total de consultas dos pacientes

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Bem Sucendida	6	9	7	6	7	7	11	2	7	4
Mal Sucendida	1	1	1	1	4	1	1	3	2	3
Neutra	8	5	7	8	4	7	3	10	6	8

Analisando a tabela 34 referentes as memórias de emoções dos agentes, observa-se que o resultado é similar às duas simulações anteriores.

A tabela 35 mostra o total de consultas de cada médico. Como era esperado, os médicos b_2 e b_3 realizaram o maior número de consultas. Já o médico b_5 teve a menor quantia de consultas, visto que ele tem uma má conduta em relação ao serviço prestado. Como era esperado, os valores mais altos das impressões, presentes na base de impressões de cada paciente (tabela 36), são as relativas aos médicos b_2 e b_3 .

Tabela 34 – Intensidade das emoções dos agentes

	<i>esperanca</i>	<i>medo</i>	<i>orgulho</i>	<i>vergonha</i>
a_1	0.88	0.4	0.83	0.4
a_2	0.98	0.4	0.99	0.4
a_3	0.97	0.4	0.94	0.4
a_4	0.54	0.5	0.58	0.5
a_5	0.55	0.5	0.7	0.5
a_6	0.84	0.5	0.82	0.5
a_7	0.55	0.5	0.5	0.5
a_8	0.2	0.8	0.2	0.8
a_9	0.2	0.92	0.2	0.9
a_{10}	0.2	0.94	0.2	0.9
b_1	x	x	0.86	0.4
b_2	x	x	0.89	0.4
b_3	x	x	0.89	0.5
b_4	x	x	0.2	0.8
b_5	x	x	0.2	0.95

Tabela 35 – Total de consultas dos médicos

	b_1	b_2	b_3	b_4	b_5
Total de consultas	30	43	41	22	14

Tabela 36 – Base de impressões dos pacientes em relação aos médicos

	b_1	b_2	b_3	b_4	b_5
a_1	0.6	0.08	0.47	0.62	-0.64
a_2	0.9	0.62	0.83	0.47	-0.29
a_3	0.53	0.54	0.53	0.43	-0.29
a_4	0.33	0.45	0.76	-0.15	0.34
a_5	0.06	0.81	0.79	-0.56	-0.15
a_6	0.36	0.38	0.63	-0.36	0.38
a_7	-0.08	0.74	0.56	-0.36	-0.22
a_8	-0.4	0.73	0.07	-0.15	-0.15
a_9	-0.5	0.84	0.14	-0.16	-0.15
a_{10}	-0.09	0.93	0.11	0.2	-0.29

Analisando as três simulações desse cenário, diferentemente do cenário 1, observa-se que a medida que o número de consultas aumenta, a

quantidade de consultas de todos os médicos também aumenta. Isso ocorre porque há um maior número de pacientes e médicos, ao mesmo tempo que o número máximo de consultas por paciente é menor se comparado ao máximo de compras realizadas por cada comprador. Essa fato torna as avaliações mais dispersas, sendo necessário um número maior de interações para que as avaliações a respeito dos médicos sejam mais precisas.